Reprinted from “Bargaining and Markets”, ISBN 0-12-528632-5,

References updated and errors corrected. Version: 2005-3-2.

Bargaining and Markets

This is a volume in
ECONOMIC THEORY, ECONOMETRICS, AND
MATHEMATICAL ECONOMICS
A series of Monographs and Textbooks

Consulting Editor: Karl Shell, Cornell University

A list of recent titles in this series appears at the end of this volume.

Bargaining and Markets

Martin J. Osborne
Department of Economics
McMaster University
Hamilton, Ontario
Canada
http://www.economics.utoronto.ca/osborne

Ariel Rubinstein
Department of Economics
Tel Aviv University
Tel Aviv, Israel
http://arielrubinstein.tau.ac.il

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
San Diego New York Boston London Sydney Tokyo Toronto

This book is printed on acid-free paper.

Academic Press, Inc.
San Diego, California 92101

United Kingdom Edition published by
Academic Press Limited
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Osborne, Martin J.
Bargaining and Markets / Martin J. Osborne and Ariel Rubinstein
p. cm.
Includes bibliographical references.
ISBN 0-12-528631-7 (alk. paper). – ISBN 0-12-528632-5 (pbk.: alk. paper)
1. Game Theory. 2. Negotiation. 3. Capitalism. I. Rubinstein, Ariel. II. Title.
HB144.073 1990
380.1–dc20 90-30644
CIP

Printed in the United States of America
90 91 92 93 9 8 7 6 5 4 3 2 1

Contents

Preface

The formal theory of bargaining originated with John Nash's work in the early 1950s. In this book we discuss two recent developments in this theory. The first uses the tool of extensive games to construct theories of bargaining in which time is modeled explicitly. The second applies the theory of bargaining to the study of decentralized markets.

We do not attempt to survey the field. Rather, we select a small number of models, each of which illustrates a key point. We take the approach that a thorough analysis of a few models is more rewarding than short discussions of many models. Some of our selections are arbitrary and could be replaced by other models that illustrate similar points.

The last section of each chapter is entitled "Notes". It usually begins by acknowledging the work on which the chapter is based. (In general we do not make acknowledgments in the text itself.) It goes on to give a brief guide to some of the related work. We should stress that this guide is not complete. We include mainly references to papers that use the model of bargaining on which most of the book is based (the bargaining game of alternating offers).

Almost always we give detailed proofs. Although this makes some of the chapters look "technical" we believe that only on understanding the proofs

is it possible to appreciate the models fully. Further, the proofs provide principles that you may find useful when constructing related models.

We use the tools of game theory throughout. Although we explain the concepts we use as we proceed, it will be useful to be familiar with the approach and basic notions of noncooperative game theory. Luce and Raiffa (1957) is a brilliant introduction to the subject. Two other recent books that present the basic ideas of noncooperative game theory are van Damme (1987) and Kreps (1990).

We have used drafts of this book for a semester-long graduate course. However, in our experience one cannot cover all the material within the time limit of such a course.

A Note on Terminology

To avoid confusion, we emphasize that we use the terms "increasing" and "nondecreasing" in the following ways. A function $f\colon \mathcal{R} \to \mathcal{R}$ for which $f(x) > f(y)$ whenever $x > y$ is *increasing*; if the first inequality is weak, the function is *nondecreasing*.

A Note on the Use of "He" and "She"

Unfortunately, the English language forces us to refer to individuals as "he" or "she". We disagree on how to handle this problem.

Ariel Rubinstein argues that we should use a "neutral" pronoun, and agrees to the use of "he", with the understanding that this refers to both men and women. Given our socio-political environment, continuous reminders of the she/he issue simply divert the reader's attention from the main issues. Language is extremely important in shaping our thinking, but in academic material it is not useful to wave it as a flag.

Martin Osborne argues that no language is "neutral". Every choice the author makes affects the reader. "He" is exclusive, and reinforces sexist attitudes, no matter how well intentioned the user. Language has a powerful impact on readers' perceptions and understanding. An author should adopt the style that is likely to have the most desirable impact on her readers' views ("the point ... is to change the world"). At present, the use of "she" for all individuals, or at least for generic individuals, would seem best to accomplish this goal.

We had to reach a compromise. When referring to specific individuals, we sometimes use "he" and sometimes "she". For example, in two-player games we treat Player 1 as female and Player 2 as male; in markets games we treat all sellers as female and all buyers as male. We use "he" for generic individuals.

Acknowledgments

The detailed comments of Ken Binmore, Jeroen Swinkels, and Eric van Damme on a draft of the book guided us to significantly improve the accuracy of the arguments and quality of the exposition. We are most grateful to them. We are grateful also to Haruo Imai, Jack Leach, Avner Shaked, Asher Wolinsky, John Wooders, and Junsen Zhang for providing valuable comments on several chapters.

Ariel Rubinstein's long and fruitful collaboration with Asher Wolinsky was the origin of many of the ideas in this book, especially those in Part 2. Asher deserves not only our gratitude but also the credit for those ideas.

Martin Osborne gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada, and thanks the Kyoto Institute of Economic Research, the Indian Statistical Institute, and the London School of Economics for their generous hospitality on visits during which he worked on this project.

Ariel Rubinstein is grateful to the London School of Economics, which was his academic home during the period in which he worked on the book.

CHAPTER 1

Introduction

1.1 Some Basic Terms

In this book we study sequential game-theoretic models of bargaining and we use them to address questions in economic theory.

1.1.1 Bargaining

Following Nash we use the term "bargaining" to refer to a situation in which (*i*) individuals ("players") have the possibility of concluding a mutually beneficial agreement, (*ii*) there is a conflict of interests about which agreement to conclude, and (*iii*) no agreement may be imposed on any individual without his approval.

A bargaining theory is an exploration of the relation between the outcome of bargaining and the characteristics of the situation. We are not concerned with questions like "what is a just agreement?", "what is a reasonable outcome for an arbitrator to decide?" or "what agreement is optimal for the society at large?" Nor do we discuss the practical issue of how to bargain effectively.

All the theories that we discuss assume that the individuals are rational, and the theories abstract from any differences in bargaining skill between individuals. We consider (in Chapter 5) the possibility that the individuals are not perfectly informed, but we maintain throughout the assumption that each individual has well-defined preferences over all relevant outcomes, and, when he has to choose between several alternatives, chooses the alternative that yields a most preferred outcome.

1.1.2 Game-Theoretic Models

Our main tool is game theory. We usually describe bargaining situations as (extensive) games. Predictions about the resolution of conflict are derived from game-theoretic solutions (variants of subgame perfect equilibrium). The analysis is intended to be precise. We do not hold the position that every claim in economic theory must be stated formally. Sometimes formal models are redundant—the arguments can be better made verbally. However, the models in this book, we believe, demonstrate the usefulness of formal models. They provide clear analyses of complex situations and lead us to a better understanding of some economic phenomena.

An interpretation of the theories in this book requires an interpretation of game theory. At several points we make comments on the interpretation of some of the notions we use, but we do not pretend to present a complete and coherent interpretation.

1.1.3 Sequentiality

Almost all the models in this book have a sequential structure: the players have to make decisions sequentially in a pre-specified order. The order reflects the procedure of bargaining (in the model in Part 1) and the procedure of trade (in the models in Part 2). The bargainers are concerned about the time at which an agreement is reached since they are impatient. The sequential structure is flexible and allows us to address a wide range of issues.

1.1.4 Economic Theory

Bargaining is a basic activity associated with trade. Even when a market is large and the traders in it take as given the environment in which they operate, there is room for bargaining when a pair of specific agents is matched. In Part 2, we study models of decentralized trade in which prices are determined by bilateral bargaining. One of the main targets of this part is to explore the circumstances under which the most basic concept of economic theory—the competitive equilibrium—is appropriate in a market in which trade is decentralized.

1.2 Outline of the Book

Part 1 contains a discussion of two theories of bargaining. We begin by studying, in **Chapter 2**, the axiomatic theory of Nash (1950a). Nash's work was the starting point for formal bargaining theory. Nash defines a "bargaining problem" to be the set of utility pairs that can be derived from possible agreements, together with a pair of utilities which is designated to be the "disagreement point". A function that assigns a single outcome to every such problem is a "bargaining solution". Nash proposes that a bargaining solution should satisfy four plausible conditions. It turns out that there is only one solution that does so, which is known as the Nash Bargaining solution. This solution has a very simple functional form, making it convenient to apply in economic models.

In **Chapter 3** we take a different tack: we impose a specific structure on bargaining and study the outcome predicted by the notion of subgame perfect equilibrium. The structure we impose is designed to keep the players as symmetric as possible. There are two players, who alternate offers. Player 1 makes an offer, which Player 2 can accept or reject; in the event of rejection, Player 2 makes a further offer, which Player 1 may accept or reject, and so on. The players have an incentive to reach an agreement because some time elapses between every offer and counteroffer—time that the players value. The game has a unique subgame perfect equilibrium, characterized by a pair of offers (x^*, y^*) with the property that Player 1 is indifferent between receiving y^* today and x^* tomorrow, and Player 2 is indifferent between receiving x^* today and y^* tomorrow. In the outcome generated by the subgame perfect equilibrium, Player 1 proposes x^*, which Player 2 accepts immediately. The simple form of this outcome lends itself to applications. We refer to the game as the bargaining game of alternating offers; it is the basic model of bargaining that we use throughout the book.

The approaches taken in Chapters 2 and 3 are very different. While the model of Chapter 2 is axiomatic, that of Chapter 3 is strategic. In the model of Chapter 2 the players' attitudes toward risk are at the forefront, while in that of Chapter 3 their attitudes to time are the driving force. Nevertheless we find in **Chapter 4** that the subgame perfect equilibrium outcome of the bargaining game of alternating offers is close to the Nash solution when the bargaining problem is defined appropriately. Given this result, each theory reinforces the other and appears to be less arbitrary.

In **Chapter 5** we turn to the analysis of bargaining in the case that one of the parties is imperfectly informed about the characteristics of his opponent. One purpose of doing so is to explain delay in reaching an agreement. We view the analysis in this chapter as preliminary because of difficulties

with the solution concept—difficulties that lie at the root of the game-theoretic modeling of situations in which players are imperfectly informed, not difficulties that are peculiar to bargaining theory. The chapter also contains a short discussion of the light the results on "mechanism design" shed on models of strategic bargaining.

Part 2 is devoted to the application of bargaining theory to the study of markets. Markets are viewed as networks of interconnected bargainers. The terms of trade between any two agents are determined by negotiation, the course of which is influenced by the agents' opportunities for trade with other partners.

One of the main targets of this literature is to understand better the circumstances under which a market is "competitive". In the theory of competitive equilibrium, the process by which the equilibrium price is reached is not modeled. One story is that there is an agency in the market that guides the price. The agency announces a price, and the individuals report the amounts they wish to demand and supply at this fixed price. If demand and supply are not equal, the agency adjusts the price. (The agency is sometimes called an "auctioneer".) This story is unpersuasive. First, we rarely observe any agency like this in actual markets. Second, it is not clear that it is possible to specify the rules used by the agency to adjust the price in such a way that it is in the interest of the individuals in the market to report truthfully their demands and supplies at any given prices.

One of our goals in studying models that probe the price-determination process is to understand the conditions (if any) under which a competitive analysis is appropriate. When it is, we consider how the primitives of the competitive model should be associated with the elements of our richer models. For example, the basic competitive model is atemporal, while the strategic models we study have a time dimension. Thus the question arises whether the demand and supply functions of the competitive model should be applied to the stock of agents in the market or to the flow of agents through the market. Also, we consider models in which the set of agents considering entering the market may be different from the set of agents who actually participate in the market. In this case we ask whether the competitive model should be applied to the demands and supplies of those in the market or of those considering entering the market.

We begin, in **Chapter 6**, by exploring models in which agents are randomly matched pairwise and conclude the agreement given by Nash's bargaining solution. We consider two basic models: one in which the number of traders in the market is steady over time (Model A), and another in which all traders enter the market at once and leave as they complete transactions (Model B). A conclusion is that the notion of competitive equilibrium fits better in the latter case. In the remainder of the book we investigate these

basic models in more detail, using strategic models of bargaining, rather than the Nash solution.

In **Chapter 7** we discuss a strategic version of Model A. Each match induces a bargaining game between the two parties. The agents are motivated to reach agreement by two factors: their own impatience and the exogenous risk that their partnership will terminate. Their utilities in the latter case depend on the equilibrium prevailing in the market; the agents take these utilities as given. We assume that the agents' behavior in the bargaining game does not depend on events in other matches. The equilibrium that we characterize does not coincide with the competitive equilibrium of the market when the demand and supply functions are those of the steady state stock of agents in the market.

In **Chapter 8** we study two strategic versions of Model B. The two models differ in the characteristics of the underlying market. In the first model, as in all other models in Part 2, each agent is either a seller or a buyer of an indivisible good. In the second model there are many divisible goods, and each agent initially holds a bundle that may contain many of these goods, as in the classical models of general equilibrium. As in Chapter 7, the agents in a matched pair may not condition their behavior on events in other matches. In both models, agents are not impatient. The models induce equilibria that correspond to the competitive equilibria of the associated markets.

In **Chapter 9** we examine how the equilibrium outcome depends on the trading procedure. For simplicity we restrict attention to markets in which there is one seller and two buyers. We are interested in the properties of the trading procedure that unleash competitive forces. We assume, in contrast to our assumption in the models of Chapters 7 and 8, that all agents are perfectly informed about all events that occur in the market (including events in matches of which they are not part). We conclude that competitive forces operate only in models in which the trading procedure allows the seller to make what is effectively a "take-it-or-leave-it" offer.

Finally, in **Chapter 10** we examine the role of the informational assumptions in the first model of Chapter 8. We find that when the agents have perfect information about all past events there are equilibria in which noncompetitive prices are sustained. Under this assumption the agents are not anonymous and thus are able to form "personal relationships".

It is not necessary to read the chapters in the order that they are presented. The dependence among them is shown in Figure 1.1. In particular, the chapters in Part 2 are largely independent of each other and do not depend on Chapters 4 and 5. Thus, if you are interested mainly in the application of bargaining theory to the study of markets, you can read Chapters 2 and 3 and then some subset of Chapters 6 through 10.

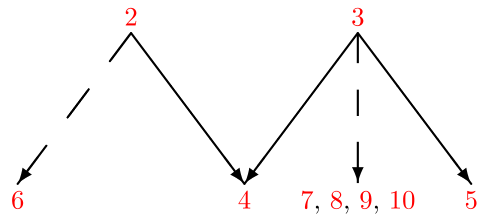


Figure 1.1 The dependence among chapters. A solid arrow indicates that the chapter above should be read before the chapter below; a dashed arrow indicates that only the main ideas from the chapter above are necessary to appreciate the chapter below.

Notes

For very compact discussions of much of the material in this book, see Wilson (1987), Bester (1989b), and Binmore, Osborne, and Rubinstein (1992). For some basic topics in bargaining theory that we do not discuss, see the following: Schelling (1960), who provides an informal discussion of the strategic elements in bargaining; Harsanyi (1977), who presents an early overview of game-theoretic models of bargaining; and Roth (1988), who discusses the large body of literature concerned with experimental tests of models of bargaining.

PART 1

Bargaining Theory

In this part we study two bargaining theories. First, in Chapter 2, we consider Nash's axiomatic model; then, in Chapter 3, we study a strategic model in which the players alternate offers. In Chapter 4 we examine the relation between the two approaches. In both models each player knows all the relevant characteristics of his opponent. In Chapter 5 we turn to the case in which the players are imperfectly informed.

CHAPTER 2

The Axiomatic Approach: Nash's Solution

2.1 Bargaining Problems

Nash (1950a) established the framework that we use to study bargaining. The set of *bargainers*—also called *players*—is N. Through most of this book we restrict attention to the case of two players: $N = \{1, 2\}$. The players either reach an *agreement* in the set A, or fail to reach agreement, in which case the *disagreement event* D occurs. Each Player $i \in N$ has a preference ordering[1] $\succsim_i$ over the set $A \cup \{D\}$. (The interpretation is that $a \succsim_i b$ if and only if Player i either prefers a to b or is indifferent between them.) The objects N, A, D, and $\succsim_i$ for each $i \in N$ define a bargaining situation.

The set A of possible agreements may take many forms. An agreement can simply be a price, or it can be a detailed contract that specifies the actions to be taken by the parties in each of many contingencies. We put no restriction directly on A. One respect in which the framework is restrictive is that it specifies a *unique* outcome if the players fail to reach agreement.

The players' attitudes toward risk play a central role in Nash's theory. We require that each player's preferences be defined on the set of lotteries over

[1] That is, a complete transitive reflexive binary relation.

possible agreements, not just on the set of agreements themselves. There is no risk explicit in a bargaining situation as we have defined it. However, uncertainty about other players' behavior, which may cause negotiation to break down, is a natural element in bargaining. Thus it is reasonable for attitudes toward risk to be part of a theory of bargaining. In fact, in Section 2.6.4 we show that there are limited possibilities for constructing an interesting axiomatic bargaining theory using as primitives only the players' preferences over agreements reached with certainty. Thus we need to enrich the model. Adding the players' attitudes toward risk is the route taken in Nash's axiomatic theory.

We assume that each player's preference ordering on the set of lotteries over possible agreements satisfies the assumptions of von Neumann and Morgenstern. Consequently, for each Player i there is a function $u_i\colon A \cup \{D\} \to \mathcal{R}$, called a *utility function*, such that one lottery is preferred to another if and only if the expected utility of the first exceeds that of the second. Such a utility function is unique only up to a positive affine transformation. Precisely, if u_i is a utility function that represents $\succsim_i$, and v_i is a utility function, then v_i represents $\succsim_i$ if and only if $v_i = \alpha u_i + \beta$ for some real numbers α and β with $\alpha > 0$.

Given the set of possible agreements, the disagreement event, and utility functions for the players' preferences, we can construct the set of all utility pairs that can be the outcome of bargaining. This is the union of the set S of all pairs $(u_1(a), u_2(a))$ for $a \in A$ and the point $d = (u_1(D), u_2(D))$. Nash takes the pair[2] $\langle S, d\rangle$ as the primitive of the problem. (Note that the same set of utility pairs could result from many different combinations of agreement sets and preferences.)

The objects of our subsequent inquiry are bargaining solutions. A bargaining solution associates with *every* bargaining situation in some class an agreement or the disagreement event. Thus, a bargaining solution does not specify an outcome for a single bargaining situation; rather, it is a function.

Formally, Nash's central definition is the following (see also Section 2.6.3).

Definition 2.1 A *bargaining problem* is a pair $\langle S, d\rangle$, where $S \subset \mathcal{R}^2$ is compact (i.e. closed and bounded) and convex, $d \in S$, and there exists $s \in S$ such that $s_i > d_i$ for $i = 1, 2$. The set of all bargaining problems is denoted $\mathcal{B}$. A *bargaining solution* is a function $f\colon \mathcal{B} \to \mathcal{R}^2$ that assigns to each bargaining problem $\langle S, d\rangle \in \mathcal{B}$ a unique element of S.

This definition restricts a bargaining problem in a number of ways. Most significantly, it eliminates the set A of agreements from the domain of

[2] Our use of angle brackets indicates that the objects enclosed are the components of a model.

discussion. Two bargaining situations that induce the same pair $\langle S, d\rangle$ are treated identically. Other theories of bargaining take A as a primitive. The assumption that the set S of feasible utility pairs is bounded means that the utilities obtainable in an outcome of bargaining are limited. The convexity assumption on S is a more significant qualitative restriction; it constrains the nature of the agreement set and utility functions. It is satisfied, for example, if A is the set of all lotteries over some underlying set of "pure" agreements (since expected utility is linear in probability). The two remaining assumptions embodied in the definition are that the players can agree to disagree ($d \in S$) and that there is some agreement preferred by both to the disagreement outcome. This last assumption ensures that the players have a mutual interest in reaching an agreement, although in general there is a conflict of interest over the particular agreement to be reached—a conflict that can be resolved by bargaining.

2.2 Nash's Axioms

Nash did not attempt to construct a model that captures all the details of any particular bargaining process; no bargaining procedure is explicit in his model. Rather, his approach is axiomatic:

> One states as axioms several properties that it would seem natural for the solution to have and then one discovers that the axioms actually determine the solution uniquely. (Nash (1953, p. 129).)

Nash imposes four axioms on a bargaining solution $f\colon \mathcal{B} \to \mathcal{R}^2$. The first formalizes the assumption that the players' preferences, not the specific utility functions that are used to represent them, are basic. We say that $\langle S', d'\rangle$ is *obtained from the bargaining problem* $\langle S, d\rangle$ *by the transformations* $s_i \mapsto \alpha_i s_i + \beta_i$ *for* $i = 1, 2$ if $d'_i = \alpha_i d_i + \beta_i$ for $i = 1, 2$, and

$$S' = \{(\alpha_1 s_1 + \beta_1, \alpha_2 s_2 + \beta_2) \in \mathcal{R}^2 \colon (s_1, s_2) \in S\}.$$

It is easy to check that if $\alpha_i > 0$ for $i = 1, 2$, then $\langle S', d'\rangle$ is itself a bargaining problem.

INV (*Invariance to Equivalent Utility Representations*) Suppose that the bargaining problem $\langle S', d'\rangle$ is obtained from $\langle S, d\rangle$ by the transformations $s_i \mapsto \alpha_i s_i + \beta_i$ for $i = 1, 2$, where $\alpha_i > 0$ for $i = 1, 2$. Then $f_i(S', d') = \alpha_i f_i(S, d) + \beta_i$ for $i = 1, 2$.

If we accept preferences, not utilities, as basic, then the two bargaining problems $\langle S, d\rangle$ and $\langle S', d'\rangle$ represent the same situation. If the utility functions u_i for $i = 1, 2$ generate the set S when applied to some set A of agreements, then the utility functions $v_i = \alpha_i u_i + \beta_i$ for $i = 1, 2$ generate

the set S' when applied to the same set A. Since v_i represents the same preferences as u_i, the physical outcome predicted by the bargaining solution should be the same for $\langle S, d\rangle$ as for $\langle S', d'\rangle$. Thus the utility outcomes should be related in the same way that the utility functions are: $f_i(S', d') = \alpha_i f_i(S, d) + \beta_i$ for $i = 1, 2$. In brief, the axiom requires that the utility outcome of bargaining co-vary with the representation of preferences, so that any physical outcome that corresponds to the solution of the problem $\langle S, d\rangle$ also corresponds to the solution of $\langle S', d'\rangle$.

Nash abstracts from any differences in "bargaining ability" between the players. If there is any asymmetry between the players then it must be captured in $\langle S, d\rangle$. If, on the other hand, the players are interchangeable, then the bargaining solution must assign the same utility to each player. Formally, the bargaining problem $\langle S, d\rangle$ is *symmetric* if $d_1 = d_2$ and $(s_1, s_2) \in S$ if and only if $(s_2, s_1) \in S$.

SYM (*Symmetry*) If the bargaining problem $\langle S, d\rangle$ is symmetric, then $f_1(S, d) = f_2(S, d)$.

The next axiom is more problematic.

IIA (*Independence of Irrelevant Alternatives*) If $\langle S, d\rangle$ and $\langle T, d\rangle$ are bargaining problems with $S \subset T$ and $f(T, d) \in S$, then $f(S, d) = f(T, d)$.

In other words, suppose that when all the alternatives in T are available, the players agree on an outcome s in the smaller set S. Then we require that the players agree on the same outcome s when only the alternatives in S are available. The idea is that in agreeing on s when they could have chosen any point in T, the players have discarded as "irrelevant" all the outcomes in T other than s. Consequently, when they are restricted to the smaller set S they should also agree on s: the solution should not depend on "irrelevant" alternatives. Note that the axiom is satisfied, in particular, by any solution that is defined to be a member of S that maximizes the value of some function.

The axiom relates to the (unmodeled) bargaining process. If the negotiators gradually eliminate outcomes as unacceptable, until just one remains, then it may be appropriate to assume IIA. On the other hand, there are procedures in which the fact that a certain agreement is available influences the outcome, even if it is not the one that is reached. Suppose, for example, that the outcome is a compromise based on the (possibly incompatible) demands of the players; such a procedure may not satisfy IIA. Without specifying the details of the bargaining process, it is hard to assess how reasonable the axiom is.

The final axiom is also problematic and, like IIA, relates to the bargaining process.

PAR (*Pareto Efficiency*) Suppose $\langle S, d\rangle$ is a bargaining problem, $s \in S$, $t \in S$, and $t_i > s_i$ for $i = 1, 2$. Then $f(S, d) \neq s$.

This requires that the players never agree on an outcome s when there is available an outcome t in which they are both better off. If they agreed on the inferior outcome s, then there would be room for "renegotiation": they could continue bargaining, the pair of utilities in the event of disagreement being s. The axiom implies that the players never disagree (since we have assumed that there is an agreement in which the utility of each Player i exceeds d_i). If we reinterpret each member of A as a pair consisting of a physical agreement and the time at which this agreement is reached, and we assume that resources are consumed by the bargaining process, then PAR implies that agreement is reached instantly.

Note that the axioms SYM and PAR restrict the behavior of the solution on single bargaining problems, while INV and IIA require the solution to exhibit some consistency across bargaining problems.

2.3 Nash's Theorem

Nash's plan of deriving a solution from some simple axioms works perfectly. He shows that there is precisely one bargaining solution satisfying the four axioms above, and this solution has a very simple form: it selects the utility pair that maximizes the product of the players' gains in utility over the disagreement outcome.

Theorem 2.2 *There is a unique bargaining solution* $f^N: \mathcal{B} \to \mathcal{R}^2$ *satisfying the axioms* INV*,* SYM*,* IIA*, and* PAR*. It is given by*

$$f^N(S, d) = \underset{(d_1,d_2)\leq(s_1,s_2)\in S}{\arg\max} (s_1 - d_1)(s_2 - d_2). \tag{2.1}$$

Proof. We proceed in a number of steps.

(a)First we verify that f^N is well defined. The set $\{s \in S : s \geq d\}$ is compact, and the function H defined by $H(s_1, s_2) = (s_1 - d_1)(s_2 - d_2)$ is continuous, so there is a solution to the maximization problem defining f^N. Further, H is strictly quasi-concave on $\{s \in S : s > d\}$, there exists $s \in S$ such that $s > d$, and S is convex, so that the maximizer is unique.

(b)Next we check that f^N satisfies the four axioms.

INV: If $\langle S', d'\rangle$ and $\langle S, d\rangle$ are as in the statement of the axiom, then $s' \in S'$ if and only if there exists $s \in S$ such that $s'_i = \alpha_i s_i + \beta_i$ for $i = 1, 2$.

For such utility pairs s and s' we have

$$(s'_1 - d'_1)(s'_2 - d'_2) = \alpha_1\alpha_2(s_1 - d_1)(s_2 - d_2).$$

Thus (s^*_1, s^*_2) maximizes $(s_1 - d_1)(s_2 - d_2)$ over S if and only if $(\alpha_1 s^*_1 + \beta_1, \alpha_2 s^*_2 + \beta_2)$ maximizes $(s'_1 - d'_1)(s'_2 - d'_2)$ over S'.

SYM: If $\langle S, d\rangle$ is symmetric and (s^*_1, s^*_2) maximizes H over S, then, since H is a symmetric function, (s^*_2, s^*_1) also maximizes H over S. Since the maximizer is unique, we have $s^*_1 = s^*_2$.

IIA: If $T \supset S$ and $s^* \in S$ maximizes H over T, then s^* also maximizes H over S.

PAR: Since H is increasing in each of its arguments, s does not maximize H over S if there exists $t \in S$ with $t_i > s_i$ for $i = 1, 2$.

(c) Finally, we show that f^N is the only bargaining solution that satisfies all four axioms. Suppose that f is a bargaining solution that satisfies the four axioms. We shall show that $f = f^N$. Let $\langle S, d\rangle$ be an arbitrary bargaining problem. We need to show that $f(S, d) = f^N(S, d)$.

Step 1. Let $f^N(S, d) = z$. Since there exists $s \in S$ such that $s_i > d_i$ for $i = 1, 2$, we have $z_i > d_i$ for $i = 1, 2$. Let $\langle S', d'\rangle$ be the bargaining problem that is obtained from $\langle S, d\rangle$ by the transformations $s_i \mapsto \alpha_i s_i + \beta_i$, which move the disagreement point to the origin and the solution $f^N(S, d)$ to the point $(1/2, 1/2)$. (That is, $\alpha_i = 1/(2(z_i - d_i))$ and $\beta_i = -d_i/(2(z_i - d_i))$, $d'_i = \alpha_i d_i + \beta_i = 0$, and $\alpha_i f^N_i(S, d) + \beta_i = \alpha_i z_i + \beta_i = 1/2$ for $i = 1, 2$.) Since both f and f^N satisfy INV we have $f_i(S', 0) = \alpha_i f_i(S, d) + \beta_i$ and $f^N_i(S', 0) = \alpha_i f^N_i(S, d) + \beta_i$ $(= 1/2)$ for $i = 1, 2$. Hence $f(S, d) = f^N(S, d)$ if and only if $f(S', 0) = f^N(S', 0)$. Since $f^N(S', 0) = (1/2, 1/2)$, it remains to show that $f(S', 0) = (1/2, 1/2)$.

Step 2. We claim that S' contains no point (s'_1, s'_2) for which $s'_1 + s'_2 > 1$. If it does, then let $(t_1, t_2) = ((1-\epsilon)(1/2) + \epsilon s'_1, (1-\epsilon)(1/2) + \epsilon s'_2)$, where $0 < \epsilon < 1$. Since S' is convex, the point (t_1, t_2) is in S'; but for ϵ small enough we have $t_1 t_2 > 1/4$ (and thus $t_i > 0$ for $i = 1, 2$), contradicting the fact that $f^N(S', 0) = (1/2, 1/2)$.

Step 3. Since S' is bounded, the result of Step 2 ensures that we can find a rectangle T that is symmetric about the 45° line and that contains S', on the boundary of which is $(1/2, 1/2)$. (See Figure 2.1.)

Step 4. By PAR and SYM we have $f(T, 0) = (1/2, 1/2)$.

Step 5. By IIA we have $f(S', 0) = f(T, 0)$, so that $f(S', 0) = (1/2, 1/2)$, completing the proof. □

Note that any bargaining solution that satisfies SYM and PAR coincides with f^N on the class of symmetric bargaining problems. The proof exploits

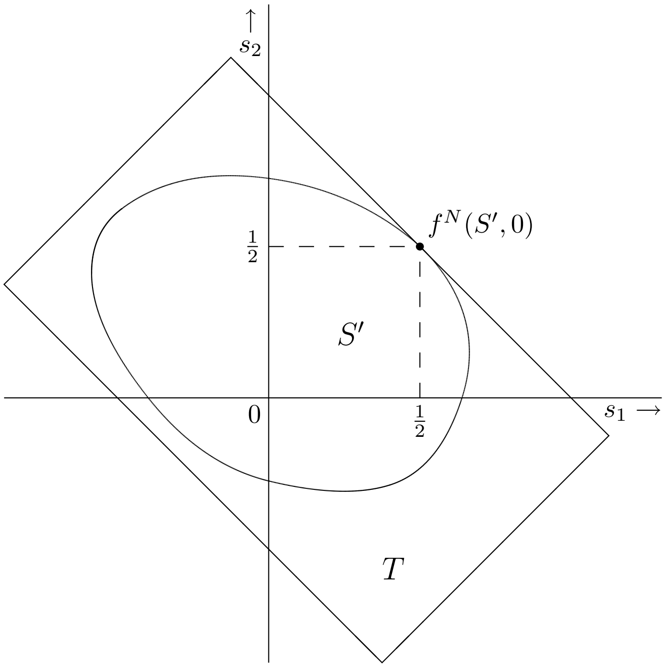


Figure 2.1 The sets S' and T in the proof of Theorem 2.2.

this fact by transforming d and $f^N(S, d)$ to points on the main diagonal, and then constructing the symmetric set T.

We refer to $f^N(S, d)$ as the *Nash solution* of the bargaining problem $\langle S, d \rangle$. It is illustrated in Figure 2.2 and can be characterized as follows. First define the *strong Pareto frontier* of S to be

$$\{s \in S\colon \text{ there is no } s' \in S \text{ with } s' \neq s \text{ and } s'_i \geq s_i \text{ for } i = 1, 2\},$$

and let $s_2 = \psi(s_1)$ be the equation of this frontier. The utility pair (s_1^*, s_2^*) is the Nash solution of $\langle S, d \rangle$ if and only if $s_2^* = \psi(s_1^*)$ and s_1^* maximizes $(s_1 - d_1)(\psi(s_1) - d_2)$. If ψ is differentiable at s_1^*, then the second condition is equivalent to $(s_2^* - d_2)/(s_1^* - d_1) = |\psi'(s_1^*)|$.

The Nash solution depends only on the preferences of the players and not on the utility representations of these preferences. However, the definition of the solution we have given is in terms of utilities. This definition is convenient in applications, but it lacks an appealing interpretation.

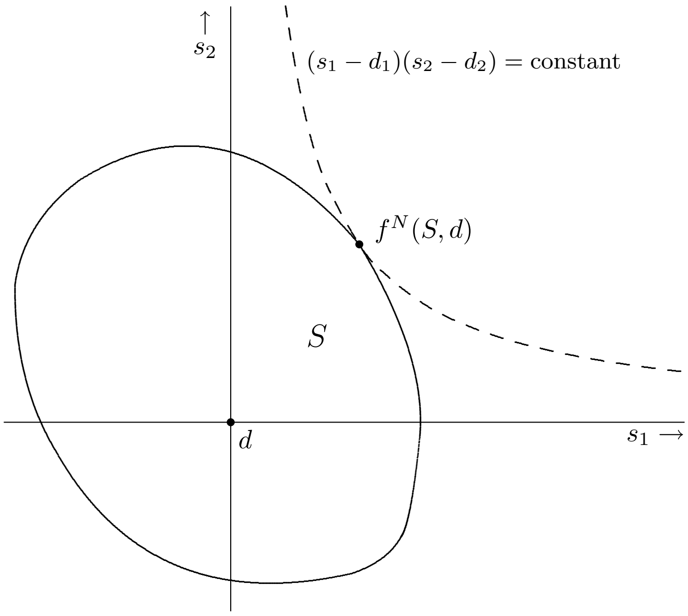


Figure 2.2 The Nash solution of the bargaining problem $\langle S, d\rangle$.

We now provide an alternative definition in terms of the players' preferences. Denote by $p \cdot a$ the lottery in which the agreement $a \in A$ is reached with probability $p \in [0, 1]$ and the disagreement event D occurs with probability $1 - p$. Let $\succsim_i$ be Player i's preference ordering over lotteries of the form $p \cdot a$, and let $\succ_i$ denote strict preference. Consider an agreement a^* with the property that for $(i, j) = (1, 2)$ and $(i, j) = (2, 1)$,

for every $a \in A$ and $p \in [0, 1]$ for which $p \cdot a \succ_i a^*$ we have $p \cdot a^* \succsim_j a$.

Any such agreement a^* has the following interpretation, which is related to that of Zeuthen (1930, Ch. IV). Assume that a^* is "on the table". If Player i is willing to object to a^* by proposing an alternative a, even if he faces the risk that with probability $1 - p$ the negotiations will break down and end with D, then Player j is willing to take the analogous risk and reject a in favor of the agreement a^*.

We now argue that any such agreement a^* induces the Nash solution of the bargaining problem. Choose utility representations u_i for $\succsim_i$ such that $u_i(D) = 0$, $i = 1, 2$. By the following argument, a^* maximizes $u_1(a)u_2(a)$.

Suppose that Player i prefers a to a^* and $u_i(a^*)/u_i(a) < u_j(a)/u_j(a^*)$. Then there exists $0 < p < 1$ such that $u_i(a^*)/u_i(a) < p < u_j(a)/u_j(a^*)$, so that $u_i(a^*) < pu_i(a)$ and $u_j(a) > pu_j(a^*)$, contradicting the definition of a^*. Hence $u_i(a^*)/u_i(a) \geq u_j(a)/u_j(a^*)$, so that $u_1(a^*)u_2(a^*) \geq u_1(a)u_2(a)$.

2.4 Applications

The simple form of the Nash solution lends itself to applications, two of which we now study.

2.4.1 Dividing a Dollar: The Role of Risk-Aversion

Two individuals can divide a dollar in any way they wish. If they fail to agree on a division, the dollar is forfeited. The individuals may, if they wish, discard some of the dollar. In terms of our model, we have

$$A = \{(a_1, a_2) \in \mathcal{R}^2 \colon a_1 + a_2 \leq 1 \text{ and } a_i \geq 0 \text{ for } i = 1, 2\}$$

(all possible divisions of the dollar), and $D = (0, 0)$ (neither player receives any payoff in the event of disagreement).

Each player is concerned only about the share of the dollar he receives: Player i prefers $a \in A$ to $b \in A$ if and only if $a_i > b_i$ ($i = 1$, 2). Thus, Player i's preferences over lotteries on A can be represented by the expected value of a utility function u_i with domain $[0, 1]$. We assume that each player is risk-averse—that is, each u_i is concave—and (without loss of generality) let $u_i(0) = 0$, for $i = 1$, 2. Then the set

$$S = \{(s_1, s_2) \in \mathcal{R}^2 \colon (s_1, s_2) = (u_1(a_1), u_2(a_2)) \text{ for some } (a_1, a_2) \in A\}$$

is compact and convex. Further, S contains $d = (u_1(0), u_2(0)) = (0, 0)$, and there is a point $s \in S$ such that $s_i > d_i$ for $i = 1, 2$. Thus $\langle S, d\rangle$ is a bargaining problem.

First, suppose that the players' preferences are the same, so that they can be represented by the same utility function. Then $\langle S, d\rangle$ is a symmetric bargaining problem. In this case, we know the Nash solution directly from SYM and PAR: it is the unique symmetric efficient utility pair $(u(1/2), u(1/2))$, which corresponds to the physical outcome in which the dollar is shared equally between the players.

If the players have different preferences, then equal division of the dollar may no longer be the agreement given by the Nash solution. Rather, the solution depends on the nature of the players' preferences. To investigate this dependence, suppose that Player 2 becomes more risk-averse. Then

his preferences, which formerly were represented by u_2, can be represented by $v_2 = h \circ u_2$, where $h: \mathcal{R} \to \mathcal{R}$ is an increasing concave function with $h(0) = 0$. (It follows that v_2 is increasing and concave, with $v_2(0) = 0$.) Player 1's preferences remain unchanged; for convenience define $v_1 = u_1$. Let $\langle S', d' \rangle$ be the bargaining problem for the new situation, in which the utility functions of the players are v_1 and v_2.

Let z_u be the solution of

$$\max_{0 \le z \le 1} u_1(z) u_2(1-z),$$

and let z_v be the solution of the corresponding problem in which v_i replaces u_i for $i = 1, 2$. Then $(u_1(z_u), u_2(1-z_u))$ is the Nash solution of $\langle S, d \rangle$, while $(v_1(z_v), v_2(1-z_v))$ is the Nash solution of $\langle S', d' \rangle$. If u_1, u_2, and h are differentiable, and $0 < z_u < 1$, then z_u is the solution of

$$\frac{u_1'(z)}{u_1(z)} = \frac{u_2'(1-z)}{u_2(1-z)}. \tag{2.2}$$

Similarly, z_v is the solution of

$$\frac{u_1'(z)}{u_1(z)} = \frac{h'\left(u_2(1-z)\right) u_2'(1-z)}{h\left(u_2(1-z)\right)}. \tag{2.3}$$

The left-hand sides of equations (2.2) and (2.3) are decreasing in z, and the right-hand sides are increasing in z. Further, since h is concave and $h(0) = 0$, we have $h'(t) \le h(t)/t$ for all t, so that the right-hand side of (2.2) is at least equal to the right-hand side of (2.3). From this we can deduce that $z_u \le z_v$, as illustrated in Figure 2.3. If $u_1 = u_2$ then we know, from the earlier argument, that $z_u = 1/2$, so that $z_v \ge 1/2$. Summarizing, we have the following.

> *If Player 2 becomes more risk-averse, then Player 1's share of the dollar in the Nash solution increases. If Player 2 is more risk-averse than Player 1, then Player 1's share of the dollar in the Nash solution exceeds* $1/2$.

Note that this result does not allow all pairs of utility functions to be compared—it applies only when one utility function is a concave function of the other.

An alternative interpretation of the problem of dividing a dollar involves the transfer of a good. Suppose that Player 1—the "seller"—holds one indivisible unit of a good, and Player 2—the "buyer"—possesses one (divisible) unit of money. The good is worthless to the seller; her utility for p units of money is $u_1(p)$, where $u_1(0) = 0$. If the buyer fails to obtain the good then his utility is zero; if he obtains the good at a price of p then

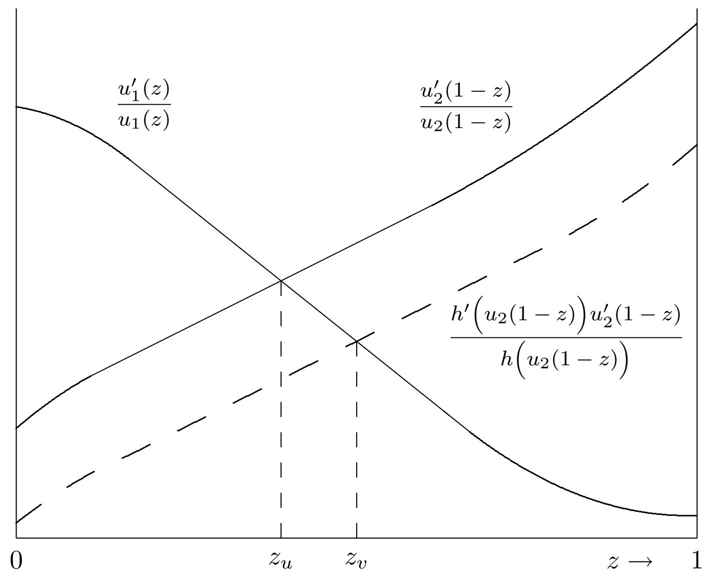


Figure 2.3 Comparative statics of the Nash solution for the problem of dividing a dollar. If the utility functions of the players are u_i ($i = 1,2$) then Player 1 receives z_u units of the dollar in the Nash solution. If Player 2 has the utility function $v_2 = h \circ u_2$, where h is increasing and concave (so that Player 2 is more risk-averse), while Player 1 retains the utility function u_1, then Player 1 receives z_v in the Nash solution.

his utility is $u_2(1 - p)$, where $u_2(0) = 0$. Both u_1 and u_2 are assumed to be concave. If the players fail to reach agreement on a sale price, then they retain their endowments. The set of agreement utility pairs for this problem is

$$S = \{(s_1, s_2) \in \mathcal{R}^2 \colon (s_1, s_2) = (u_1(p), u_2(1-p)) \text{ for some } 0 \leq p \leq 1\}$$

and the disagreement point is $d = (0, 0)$, so that the problem is formally identical to that of dividing a dollar.

2.4.2 Negotiating Wages

In the example above, an agreement specifies the amount of money to be received by each party. In other cases, an agreement may be very complex. In the context of negotiation between a firm and a labor union, for example, an agreement may specify a stream of future wages, benefits, and employment levels.

To illustrate a relatively simple case, consider a firm and a union negotiating a wage-employment package. Suppose that the union represents L workers, each of whom can obtain a wage of w_0 outside the firm. If the firm hires ℓ workers, then it produces $f(\ell)$ units of output. We assume that f is strictly concave, $f(0) = 0$, and $f(\ell) > \ell w_0$ for some ℓ, and normalize the price of output to be one. An agreement is a wage-employment pair (w, ℓ). The von Neumann–Morgenstern utility of the firm for the agreement (w, ℓ) is its profit $f(\ell) - \ell w$, while that of the union is the total amount of money $\ell w + (L - \ell)w_0$ received by its members. (This is one of a number of possible objectives for the union.) We restrict agreements to be pairs (w, ℓ) in which the profit of the firm is nonnegative ($w \leq f(\ell)/\ell$) and the wage is at least w_0. Thus the set of utility pairs that can result from agreement is

$$S = \{(f(\ell) - \ell w, \ell w + (L - \ell)w_0) \colon f(\ell) \geq \ell w,\ 0 \leq \ell \leq L \text{ and } w \geq w_0\}.$$

If the two parties fail to agree, then the firm obtains a profit of zero (since $f(0) = 0$) and the union receives Lw_0, so that the disagreement utility pair is $d = (0, Lw_0)$.

Each pair of utilities takes the form $(f(\ell) - \ell w, \ell w + (L - \ell)w_0)$, where $w_0 \leq w \leq f(\ell)/\ell$. Let ℓ^* be the unique maximizer of $f(\ell) + (L - \ell)w_0$. Then the set of utility pairs that can be attained in an agreement is

$$S = \{(s_1, s_2) \in \mathcal{R}^2 \colon s_1 + s_2 \leq f(\ell^*) + (L - \ell^*)w_0,\ s_1 \geq 0, \text{ and } s_2 \geq Lw_0\}.$$

This is a compact convex set, which contains the disagreement point $d = (0, Lw_0)$ in its interior. Thus $\langle S, d \rangle$ is a bargaining problem.

Given that the Nash solution is efficient (i.e. it is on the Pareto frontier of S), the size of the labor force it predicts is ℓ^*, which maximizes the profit $f(\ell) - \ell w_0$. To find the wage it predicts, note that the difference between the union's payoff at the agreement (w, ℓ) and its disagreement payoff is $\ell w + (L - \ell)w_0 - Lw_0 = \ell(w - w_0)$. Thus the predicted wage is

$$\arg\max_{w \geq w_0} (f(\ell^*) - \ell^* w)\ell^*(w - w_0).$$

This is $w^* = (w_0 + f(\ell^*)/\ell^*)/2$: the average of the outside wage and the average product of labor.

2.5 Is Any Axiom Superfluous?

We have shown that Nash's four axioms uniquely define a bargaining solution, but have not ruled out the possibility that some subset of the axioms is enough to determine the solution uniquely. We now show that none of the

axioms is superfluous. We do so by exhibiting, for each axiom, a solution that satisfies the remaining three axioms and is different from Nash's.

INV: Let $g: \mathcal{R}^2_+ \to \mathcal{R}$ be increasing and strictly quasi-concave, and suppose that each of its contours $g(x_1, x_2) = c$ has slope -1 when $x_1 = x_2$. Consider the bargaining solution that assigns to each bargaining problem $\langle S, d \rangle$ the (unique) maximizer of $g(s_1 - d_1, s_2 - d_2)$ over $\{s \in S : s \geq d\}$. This solution satisfies PAR and IIA (since it is the maximizer of an increasing function) and also SYM (by the condition on the slope of its contours). To show that this solution differs from that of Nash, let $g(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2}$ and consider the bargaining problem $\langle S, d \rangle$ in which $d = (0, 0)$ and S is the convex hull[3] of the points $(0, 0)$, $(1, 0)$, and $(0, 2)$. The maximizer of g for this problem is $(s_1, s_2) = (1/3, 4/3)$, while its Nash solution is $(1/2, 1)$.

Another solution that satisfies PAR, IIA, and SYM, and differs from the Nash bargaining solution, is given by that maximizer of $s_1 + s_2$ over $\{s \in S : s \geq d\}$ that is closest to the line with slope 1 through d. This solution is appealing since it simply maximizes the *sum* (rather than the product, as in Nash) of the excesses of the players' utilities over their disagreement utilities.

SYM: For each $\alpha \in (0, 1)$ consider the solution f^α that assigns to $\langle S, d \rangle$ the utility pair

$$\arg\max_{(d_1,d_2) \leq (s_1,s_2) \in S} (s_1 - d_1)^\alpha (s_2 - d_2)^{1-\alpha}. \tag{2.4}$$

The family of solutions $\{f^\alpha\}_{\alpha \in (0,1)}$ is known as the family of *asymmetric Nash solutions*. For the problem $\langle S, d \rangle$ in which $d = (0, 0)$ and S is the convex hull of $(0, 0)$, $(1, 0)$, and $(0, 1)$, we have $f^\alpha(S, d) = (\alpha, 1 - \alpha)$, which, when $\alpha \neq 1/2$, is different from the Nash solution of $\langle S, d \rangle$. Every solution f^α satisfies INV, IIA, and PAR by arguments exactly like those used for the Nash solution.

IIA: For any bargaining problem $\langle S, d \rangle$, let $\bar{s}_i$ be the maximum utility Player i gets in $\{s \in S : s \geq d\}$, for $i = 1, 2$. Consider the solution $f^{KS}(S, d)$ that assigns to $\langle S, d \rangle$ the maximal member of S on the line joining d and $(\bar{s}_1, \bar{s}_2)$ (see Figure 2.4).

For the bargaining problem in which $d = (0, 0)$ and S is the convex hull of $(0, 0)$, $(1, 0)$, $(1/2, 1/2)$, and $(0, 1/2)$, we have $f^{KS}(S, d) = (2/3, 1/3)$, different from the utility pair $(1/2, 1/2)$ predicted by the Nash solution. It is immediate that the solution satisfies SYM and PAR; it is straightforward

[3] The convex hull of a finite set of points is the smallest convex set (a polyhedron) containing the points.

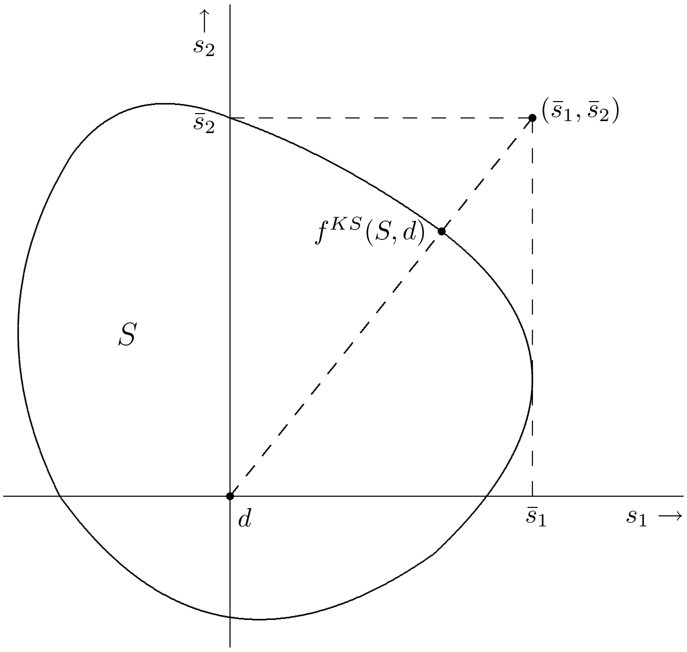


Figure 2.4 The Kalai–Smorodinsky solution f^{KS}.

to argue that it also satisfies INV. This solution is known as the Kalai–Smorodinsky solution.

PAR: Consider the solution f^d defined by $f^d(S,d) = d$. This solution satisfies INV, SYM, and IIA and is different from the Nash solution.

For each of the four axioms, we have described a solution different from Nash's that satisfies the remaining three axioms. Some of these solutions have interesting axiomatizations.

Say that a bargaining solution f satisfies *strong individual rationality* (SIR) if $f(S,d) > d$ for every bargaining problem $\langle S,d \rangle$. Then a solution satisfies INV, PAR, IIA, and SIR if and only if it is an asymmetric Nash solution (i.e. of the form f^α for some $\alpha \in (0,1)$).

The Kalai–Smorodinsky solution f^{KS} is the only solution satisfying INV, SYM, PAR, and a "monotonicity" axiom. The last axiom requires that if $S \subset T$ and if, for $i = 1, 2$, the maximum utilities that Player i can obtain in

$\{s \in S: s \geq d\}$ and $\{s \in T: s \geq d\}$ are the same, then each player receives at least as much utility in the solution of $\langle T, d\rangle$ as in the solution of $\langle S, d\rangle$.

Finally, the solution f^d, which assigns to every bargaining problem the disagreement point, is the only solution other than Nash's that satisfies INV, SYM, IIA, and the condition that the solution give each player at least his disagreement utility. Thus PAR can be replaced by SIR in the characterization of the Nash solution.

2.6 Extensions of the Theory

2.6.1 More Than Two Players

All our arguments concerning Nash's solution can be extended to situations in which there are more than two players. If there are n players then a bargaining problem is a pair $\langle S, d\rangle$, in which S is a compact convex subset of $\mathcal{R}^n$, $d \in S$, and there exists $s \in S$ such that $s_i > d_i$ for $i = 1, \ldots, n$. The four axioms INV, SYM, IIA, and PAR can be extended to apply to bargaining problems with n players, and it can be shown that the unique bargaining solution that satisfies the axioms is the function that associates with each problem $\langle S, d\rangle$ the vector of utilities

$$\arg\max_{d \leq s \in S} \prod_{i=1}^{n} (s_i - d_i).$$

2.6.2 An Alternative Interpretation of Utility

So far we have interpreted the elements of S as utility pairs derived from the players' preferences over lotteries on the set of physical agreements. We have assumed that these preferences satisfy the assumptions of von Neumann and Morgenstern, so that the utility functions that represent them are unique only up to an affine transformation. Under this assumption, the axiom INV is appropriate.

There are alternative interpretations in which risk does not enter explicitly. To make sense of Nash's theory we require only that the utilities represent the underlying preferences uniquely up to an affine transformation. An alternative to starting with a player's preferences over lotteries is to adopt as a primitive the player's preference over the set of agreements, combined with his answers to all possible questions of the form "Do you prefer a to a' more than you prefer b to b'?" Under additional assumptions the player's preferences can be represented by a utility function that is unique up to an affine transformation (see Krantz, Luce, Suppes, and Tversky (1971, Ch. 4)).

We previously interpreted the results in Section 2.4.1 as showing the effect on the Nash solution of changes in the degree of risk-aversion of the players. Under the alternative interpretation of utility presented here, these results show the effect of changes in the rate at which players' marginal utilities decline. More precisely, the results are the following. If the marginal utility of Player $2'$ declines more rapidly than that of Player 2, then Player 1's share of the dollar in the Nash solution of the divide-the-dollar game is larger when his opponent is Player $2'$ than when it is Player 2. If Player 2's marginal utility declines more rapidly than that of Player 1, then Player 1's share of the dollar in the Nash solution exceeds 1/2.

2.6.3 An Alternative Definition of a Bargaining Problem

A bargaining problem, as we have defined it so far, consists of a compact convex set $S \subset \mathcal{R}^2$ and an element $d \in S$. However, the Nash solution of $\langle S, d\rangle$ depends only on d and the Pareto frontier of S. The compactness and convexity of S are important only insofar as they ensure that the Pareto frontier of S is well defined and concave. Rather than starting with the set S, we could have imposed our axioms on a problem defined by a nonincreasing concave function (and a disagreement point d). In the following chapters it will frequently be more natural to define a problem in this way. Thus, we sometimes subsequently use the term *bargaining problem* to refer to a pair $\langle S, d\rangle$, where S is the graph of a nonincreasing concave real-valued function defined on a closed interval of $\mathcal{R}$, $d \in \mathcal{R}^2$, and there exists $s \in S$ such that $s_i > d_i$ for $i = 1, 2$.

2.6.4 Can a Solution Be Based on Ordinal Preferences?

Utility functions are not present in Nash's formal model of a bargaining problem, which consists solely of a set S and a point $d \in S$. Nevertheless, the cardinality of the players' utility functions (a consequence, for example, of the fact that they are derived from preferences over lotteries that satisfy the assumptions of von Neumann and Morgenstern) plays a major role in the arguments we have made. In particular, the appeal of the invariance axiom INV derives from the fact that utility functions that differ by a positive affine transformation represent the same preferences.

It is natural to ask whether a theory of bargaining can be based purely on ordinal information about preferences. Suppose we are given the set A of physical agreements, the disagreement outcome $D \in A$, and the (ordinal) preferences of each player over A. Can we construct a theory of bargaining based on only these elements? It is possible to do so if we retreat from Nash's basic assumption that the physical characteristics of members of A

do not matter. However, we face a difficulty if we want the theory to select an outcome that depends only on the players' preferences for elements of A. In order to construct such a theory, we must be able to describe the outcome solely in terms of preferences. For example, the theory could predict the outcome that Player 1 most prefers. Or it could predict the outcome that Player 2 most prefers among those that Player 1 finds indifferent to the disagreement outcome. Neither of these appears to be a plausible outcome of bargaining. Further, intuition suggests that the set of outcomes that can be described in terms of preferences may be very small and not contain *any* "reasonable" outcome.

We can make a precise argument as follows. If we are to succeed in constructing a bargaining theory that depends only on the data $\langle A, D, \succsim_1, \succsim_2\rangle$, then our bargaining solution F must satisfy the following condition (an analog of INV). Let $\langle A, D, \succsim_1, \succsim_2\rangle$ and $\langle A', D', \succsim_1', \succsim_2'\rangle$ be two bargaining problems, and let $T\colon A \to A'$ be a one-to-one function with $T(A) = A'$. Suppose that T preserves the preference orderings (i.e. $T(a) \succsim_i' T(b)$ if and only if $a \succsim_i b$ for $i = 1$, 2) and satisfies $T(D) = D'$. Then T must transform the solution of the first problem into the solution of the second problem: $T\left(F(A, D, \succsim_1, \succsim_2)\right) = F(A', D', \succsim_1', \succsim_2')$. In particular, if $\langle A, D, \succsim_1, \succsim_2\rangle = \langle A', D', \succsim_1', \succsim_2'\rangle$, then T must map the solution into itself.

Now suppose that the set of agreements is

$$A = \{(a_1, a_2) \in \mathcal{R}^2 \colon a_1 + a_2 \leq 1 \text{ and } a_i \geq 0 \text{ for } i = 1, 2\},$$

the disagreement outcome is $D = (0, 0)$, and the preference ordering of each Player $i = 1$, 2 is defined by $(a_1, a_2) \succsim_i (b_1, b_2)$ if and only if $a_i \geq b_i$. Define $T\colon A \to A$ by $T(a_1, a_2) = (2a_1/(1 + a_1), a_2/(2 - a_2))$. This maps A onto itself, satisfies $T(0, 0) = (0, 0)$, and preserves the preference orderings. However, the only points (a_1, a_2) for which $T(a_1, a_2) = (a_1, a_2)$ are $(0, 0)$, $(1, 0)$, and $(0, 1)$. Thus a bargaining theory based solely on the information $\langle A, D, \succsim_1, \succsim_2\rangle$ must assign one of these three outcomes to be the bargaining solution. Since none of them is attractive as a solution, we conclude that no nontrivial theory of bargaining that includes this problem can be based on ordinal preferences.

Note that if the number of alternatives is finite, then the arguments we have made no longer apply: in this case, many alternatives may be given an ordinal description. Note further that our argument is limited to the case of two players. We have not established that a nontrivial theory of bargaining based on ordinal preferences is impossible when we restrict attention to problems in which there are three or more players. Indeed, in this case Shubik (1982, pp. 96–98) exhibits such a theory.

2.6.5 Nash's "Variable Threat" Model

In Nash's axiomatic model the point d, which is interpreted as the outcome in the event that the players fail to reach agreement, is fixed. Nash (1953) extended his theory to encompass situations in which the players can influence this outcome. The primitive of this later model is a two-person strategic game, which we denote G, in which each player has finitely many pure strategies. Let P_i be Player i's set of pure strategies, let Σ_i be his set of mixed strategies (i.e. probability distributions over pure strategies), and let $H_i: \Sigma_1 \times \Sigma_2 \to \mathcal{R}$ be his payoff function.

The players begin by simultaneously selecting mixed strategies in G. These strategies are interpreted as the actions the players are *bound* to take if they fail to reach agreement; we refer to them as *threats*. The players must carry out their threats in case of disagreement even when the pair of threats is not a Nash equilibrium of G. Once the threats have been chosen, the agreement that is reached is given by the Nash solution of the bargaining problem in which the set of possible agreements is the set of probability distributions over $P_1 \times P_2$, and the disagreement point is the pair of payoffs in G in the event the threats are carried out. Given the threat of Player j, Player i's payoff in the Nash solution is affected by his own threat; each player chooses his threat to maximize his payoff, given the threat of the other player.

More precisely, let S be the (convex and compact) set of pairs of payoffs to probability distributions over $P_1 \times P_2$, and define the function $g: S \to S$ by $g(d) = f^N(S, d)$, where f^N is the Nash solution function. Nash's *threat game* is the game G^* in which Player i's pure strategy set is Σ_i and his payoff to the strategy pair (σ_1, σ_2) is $g_i(H(\sigma_1, \sigma_2))$, where $H(\sigma_1, \sigma_2) = (H_1(\sigma_1, \sigma_2), H_2(\sigma_1, \sigma_2))$. The game G^* has a Nash equilibrium (which is sometimes referred to as a pair of *optimal threats*). This follows from a standard result on the existence of Nash equilibrium, given that $g_i \circ H$ is continuous and quasi-concave in σ_i for each given value of σ_j. Since G^* is strictly competitive (i.e. $g_1(H(\sigma_1, \sigma_2)) > g_1(H(\sigma_1', \sigma_2'))$ if and only if $g_2(H(\sigma_1, \sigma_2)) < g_2(H(\sigma_1', \sigma_2'))$), each player's equilibrium strategy guarantees him his equilibrium payoff.

Notes

The main body of this chapter (Sections 2.1, 2.2, and 2.3) is based on Nash's seminal paper (1950a). We strongly urge you to read this paper.

The effect of the players' risk-aversion on the agreement predicted by the Nash solution (considered in Section 2.4.1) is explored by Kihlstrom, Roth, and Schmeidler (1981). The analysis of wage negotiation in Section 2.4.2

appears in McDonald and Solow (1981). Harsanyi and Selten (1972) study the asymmetric Nash solutions f^α described in Section 2.5. Precise axiomatizations of these solutions, along the lines of Nash's Theorem, are given by Kalai (1977) and Roth (1979, p. 16). (The one stated in the text is Kalai's.) The axiomatization of f^{KS} is due to Kalai and Smorodinsky (1975). Roth (1977) shows that PAR may be replaced by SIR in the characterization of the Nash solution. The n-player Nash solution defined in Section 2.6.1 is studied, for example, by Roth (1979). The discussion in Section 2.6.4 of a bargaining theory that uses only ordinal information about preferences is based on Shapley (1969). The model in Section 2.6.5 is due to Nash (1953); for further discussion see Owen (1982, Section VII.3) and Hart (1979, Ch. I).

The literature on the axiomatic approach to bargaining is surveyed by Roth (1979) and Thomson (forthcoming), and, more compactly, by Kalai (1985).

CHAPTER 3

The Strategic Approach: A Model of Alternating Offers

3.1 The Strategic Approach

In the axiomatic approach, the outcome of bargaining is defined by a list of properties that it is required to satisfy. In the strategic approach, the outcome is an equilibrium of an explicit model of the bargaining process. In Nash's words,

> one makes the players' steps of negotiation ... moves in the non-cooperative model. Of course, one cannot represent all possible bargaining devices as moves in the non-cooperative game. The negotiation process must be formalized and restricted, but in such a way that each participant is still able to utilize all the essential strengths of his position. (Nash (1953, p. 129).)

Any strategic model embodies a detailed description of a bargaining procedure. Since we observe a bewildering variety of such procedures, we are faced with the difficult task of formulating a tractable model that expresses the main influences on the outcome. A complex model that imposes little structure on the negotiation is unlikely to yield definite results; a simple model may omit a key element in the determination of the settlement. With this tradeoff in mind, we construct in this chapter a model that focuses on just one salient feature of bargaining: the participants' attitudes to delay.

3.2 The Structure of Bargaining

The situation we model is the following. Two players bargain over a "pie" of size 1. An agreement is a pair (x_1, x_2), in which x_i is Player i's share of the pie. The set of possible agreements is

$$X = \{(x_1, x_2) \in \mathcal{R}^2 \colon x_1 + x_2 = 1 \text{ and } x_i \geq 0 \text{ for } i = 1, 2\}.$$

The players' preferences over X are diametrically opposed. Each player is concerned only about the share of the pie that he receives, and prefers to receive more rather than less. That is, Player i prefers $x \in X$ to $y \in X$ if and only if $x_i > y_i$. Note that X is the set of *agreements*, not the set of *utility pairs*; we are *not* restricting attention to the case in which the latter is a line segment.

This is a simple setting, but it is rich enough to include the examples discussed in Section 2.4. In the case of bargaining over the division of a dollar, we interpret x_i as the amount that Player i receives. In the case of negotiating the sale price of an indivisible good, x_1 is the price the buyer pays to the seller. In the model of wage negotiation, x_1 is the profit of the firm.

The bargaining procedure is as follows. The players can take actions only at times in the (infinite) set $T = \{0, 1, 2, \ldots\}$. In each period $t \in T$ one of the players, say i, proposes an agreement (a member of X), and the other player (j) either accepts the offer (chooses Y) or rejects it (chooses N). If the offer is accepted, then the bargaining ends, and the agreement is implemented. If the offer is rejected, then the play passes to period $t + 1$; in this period Player j proposes an agreement, which Player i may accept or reject. The game continues in this manner; whenever an offer is rejected, play passes to the next period, in which it is the rejecting player's turn to propose an agreement. There is no limit on the number of periods. Once an offer has been rejected, it is void; the player who made the offer is free to propose any agreement in the future, and there is no restriction on what he may accept or reject. At all times, each player knows all his previous moves and all those of the other player.

We model this procedure as an extensive game. For convenience, we give the game an explicit time structure. The first two periods of the game are shown in Figure 3.1. Play begins at the top of the tree, and time starts at period 0. The number beside each node indicates the player whose turn it is to move there. Thus Player 1 is the first to move; she has a continuum of choices (indicated by the triangle attached to her decision node). This continuum corresponds to the agreements (members of X) that Player 1 can propose. Each possible proposal leads to a decision

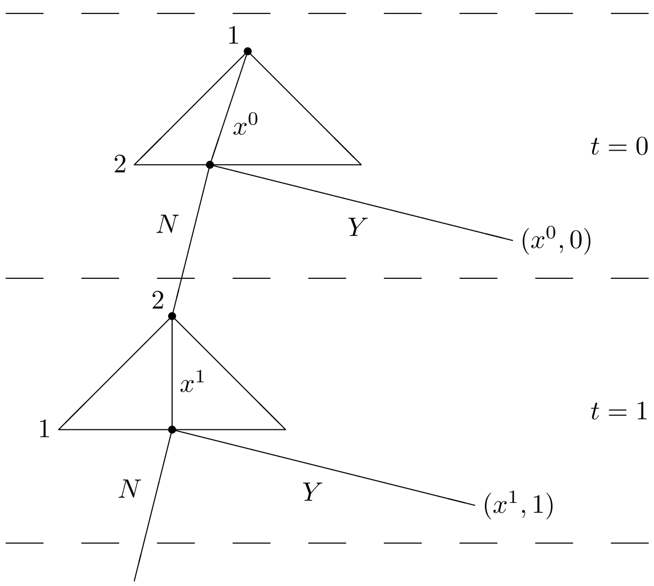


Figure 3.1 The first two periods of a bargaining game of alternating offers. The number beside each node is the player who takes an action there. The branch labelled x^0 represents a "typical" offer of Player 1 out of the continuum available in period 0. The labels Y and N refer to the actions "accept" and "reject".

node for Player 2, at which he accepts (Y) or rejects (N) the proposal. One such node, corresponding to the proposal x^0 is indicated. If Player 2 agrees (the right-hand branch), then the game ends; the label $(x^0, 0)$ indicates that the agreement x^0 is reached in period 0. If Player 2 rejects Player 1's offer (the left-hand branch), then play passes to period 1, when it is Player 2's turn to make an offer. A typical offer of Player 2 is x^1; for each such offer, Player 1 says either Y or N. If Player 1 chooses Y, then the game ends with the outcome $(x^1, 1)$; if she chooses N then the game continues—Player 1 makes a further offer, Player 2 responds, and so on.

Note that the tree is infinite in two respects. First, at any node where a player makes an offer, there is a *continuum*, rather than a finite number of choices. Consequently, it is not possible to show in the diagram *every* subsequent node of the other player; we have selected one "typical" choice at each such point. Second, the tree contains unboundedly long paths, in which *all* offers are rejected, so that a terminal node is never reached. We

assume that every such path leads to the *same* outcome, which we denote D ("disagreement"). Note also that the roles of the players are almost symmetric; the only asymmetry is that Player 1 is the first to make an offer.

We have attached to the terminal nodes (those at which an agreement is concluded) labels of the form (x, t), giving the nature of the agreement and the time at which it is reached, rather than labeling these nodes with payoffs. Also we have assumed that all the infinite paths (at which an agreement is never reached) lead to the *same* outcome D. In order to analyze the players' choices we have to specify their preferences over these outcomes. But before we do so (in the next section), note that in defining an outcome to be either a pair (x, t) or D, we have made a restrictive assumption about these preferences. For any period $t \geq 1$, *many* paths through the tree lead to a terminal node with label (x, t), since we assign this outcome whenever the players agree to x at time t, *regardless* of the previous path of rejected offers. (The diagram in Figure 3.1 obscures this, since it contains only one "typical" offer at each stage.) Thus we assume that players care only about the nature of the agreement and the time at which it is reached, *not* about the sequence of offers and counteroffers that leads to the agreement.[1] In particular, no player *regrets* having made an offer that was rejected. Similarly, we assume that the players are indifferent about the sequence of offers and rejections that leads to disagreement.

Finally, note that the structure of the game is different from that of a repeated game. The structure of the tree is repetitive, but once a player accepts an offer, the game ceases to be repeated.

3.3 Preferences

3.3.1 Assumptions

In Nash's axiomatic approach, the players' preferences over the physical outcomes are augmented by their attitudes toward risk; we saw (Section 2.6) that preferences over the physical outcomes alone may not be sufficient to determine a solution. Here, the structure of the game requires us to include in the description of the players' preferences their attitudes toward agreements reached at various points in time. These *time preferences* are the driving force of the model.

In order to complete our description of the game we need to specify the players' preferences. We assume that each player $i = 1, 2$ has a complete

[1] In using the word "outcome" for a pair (x, t) we are deviating slightly from normal usage, in which "outcome" refers to a path through the tree.

transitive reflexive preference ordering[2] $\succsim_i$ over the set $(X \times T) \cup \{D\}$ of *outcomes*.

Definition 3.1 A *bargaining game of alternating offers* is an extensive game with the structure defined in Section 3.2, in which each player's preference ordering $\succsim_i$ over $(X \times T) \cup \{D\}$ is complete, transitive, and reflexive.

In the main analysis of this chapter we impose a number of conditions on the players' preference orderings. These conditions are weak enough to allow a wide variety of preferences. In particular, preferences over $X \times T$ for Player i that are represented by the function $\delta_i^t u_i(x_i)$ for any $0 < \delta_i < 1$ and any increasing concave function u_i are allowed. Specifically, our assumptions are the following. First, the least-preferred outcome is D.

A1 (*Disagreement is the worst outcome*) For every $(x, t) \in X \times T$ we have $(x, t) \succsim_i D$.

The remaining conditions concern the behavior of $\succsim_i$ on $X \times T$. First, we require that among agreements reached in the same period, Player i prefers larger values of x_i and prefers to obtain any given share of the pie sooner rather than later.

A2 (*Pie is desirable*) For any $t \in T$, $x \in X$, and $y \in X$ we have $(x, t) \succ_i (y, t)$ if and only if $x_i > y_i$.

A3 (*Time is valuable*) For any $t \in T$, $s \in T$, and $x \in X$ we have $(x, t) \succsim_i (x, s)$ if $t < s$, with strict preference if $x_i > 0$.

Next we assume that Player i's preference ordering is continuous.

A4 (*Continuity*) Let $\{(x_n, t)\}_{n=1}^{\infty}$ and $\{(y_n, s)\}_{n=1}^{\infty}$ be sequences of members of $X \times T$ for which $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. Then $(x, t) \succsim_i (y, s)$ whenever $(x_n, t) \succsim_i (y_n, s)$ for all n.

The ordering $\succsim_i$ satisfies assumptions A2 through A4 if and only if i's preferences over $X \times T$ can be represented by a continuous utility function $U_i\colon [0, 1] \times T \to \mathcal{R}$ that is increasing in its first argument (the share of the pie received by i), and decreasing in its second argument (the period of receipt) when the first argument is positive. (See Fishburn and Rubinstein (1982, Theorem 1).[3])

---

[2]Following convention, we write $z \sim_i z'$ if $z \succsim_i z'$ and $z' \succsim_i z$, and say that z and z' are *indifferent* for Player i; we write $z \succ_i z'$ if it not true that $z' \succsim_i z$.

[3]Fishburn and Rubinstein assume, in addition to A3, that $(x, t) \sim_i (x, s)$ for all $t \in T$ and $s \in T$ whenever $x_i = 0$. However, their proof can easily be modified to deal with the case in which the weaker condition in A3 is satisfied.

The next assumption greatly simplifies the structure of preferences. It requires that the preference between (x, t) and (y, s) depend only on x, y, and the *difference* $s - t$. Thus, for example, it implies that if $(x, 1) \sim_i (y, 2)$ then $(x, 4) \sim_i (y, 5)$.

A5 (*Stationarity*) For any $t \in T$, $x \in X$, and $y \in X$ we have $(x, t) \succ_i (y, t + 1)$ if and only if $(x, 0) \succ_i (y, 1)$.

If the ordering $\succeq_i$ satisfies A5 in addition to A2 through A4 then there is a utility function U_i representing i's preferences over $X \times T$ that has a specific form: for every $\delta \in (0, 1)$ there is a continuous increasing function $u_i \colon [0, 1] \to \mathcal{R}$ such that $U_i(x_i, t) = \delta^t u_i(x_i)$. (See Fishburn and Rubinstein (1982, Theorem 2).[4]) Note that for *every* value of δ we can find a suitable function u_i; the value of δ is not determined by the preferences. Note also that the function u_i is not necessarily concave.

To facilitate the subsequent analysis, it is convenient to introduce some additional notation. For any outcome (x, t), it follows from A2 through A4 that either there is a unique $y \in X$ such that Player i is indifferent between (x, t) and $(y, 0)$ (in which case A3 implies that if $x_i > 0$ and $t \geq 1$ then $y_i < x_i$), or *every* outcome $(y, 0)$ (including that in which $y_i = 0$) is preferred by i to (x, t). Define $v_i \colon [0, 1] \times T \to [0, 1]$ for $i = 1, 2$ as follows:

$$v_i(x_i, t) = \begin{cases} y_i & \text{if } (y, 0) \sim_i (x, t) \\ 0 & \text{if } (y, 0) \succ_i (x, t) \text{ for all } y \in X. \end{cases} \tag{3.1}$$

The analysis may be simplified by making the more restrictive assumption that for all (x, t) and for $i = 1, 2$ there exists y such that $(y, 0) \sim_i (x, t)$. This restriction rules out some interesting cases, and therefore we do not impose it. However, to make a first reading of the text easier we suggest that you adopt this assumption.

It follows from (3.1) that if $v_i(x_i, t) > 0$ then Player i is indifferent between receiving $v_i(x_i, t)$ in period 0 and x_i in period t. We slightly abuse the terminology and refer to $v_i(x_i, t)$ as the *present value of* (x, t) *for Player* i even when $v_i(x_i, t) = 0$. Note that

$$(y, 0) \succeq_i (x, t) \text{ whenever } y_i = v_i(x_i, t) \tag{3.2}$$

and $(y, t) \succ_i (x, s)$ whenever $v_i(y_i, t) > v_i(x_i, s)$.

If the preference ordering $\succeq_i$ satisfies assumptions A2 through A4, then for each $t \in T$ the function $v_i(\cdot, t)$ is continuous, nondecreasing, and increasing whenever $v_i(x_i, t) > 0$; further, we have $v_i(x_i, t) \leq x_i$ for every $(x, t) \in X \times T$, and $v_i(x_i, t) < x_i$ whenever $x_i > 0$ and $t \geq 1$. Under A5 we have $v_i\left(v_i(x_i, 1), 1\right) = v_i(x_i, 2)$ for any $x \in X$. An example of the functions $v_1(\cdot, 1)$ and $v_2(\cdot, 1)$ is shown in Figure 3.2.

[4]The comment in the previous footnote applies.

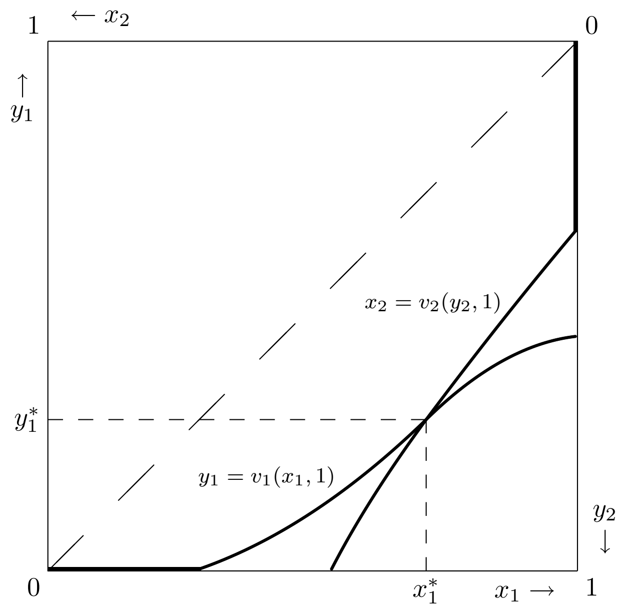


Figure 3.2 The functions $v_1(\cdot, 1)$ and $v_2(\cdot, 1)$. The origin for the graph of $v_1(\cdot, 1)$ is the lower left corner of the box; the origin for the graph of $v_2(\cdot, 1)$ is the upper right corner.

Under assumption A3 any given amount is worth less the later it is received. The final condition we impose on preferences is that the loss to delay associated with any given amount is an increasing function of the amount.

A6 (*Increasing loss to delay*) The difference $x_i - v_i(x_i, 1)$ is an increasing function of x_i.

Under this assumption the graph of each function $v_i(\cdot, 1)$ in Figure 3.2 has a slope (relative to its origin) of less than 1 everywhere. The assumption also restricts the character of the function u_i in any representation $\delta^t u_i(x_i)$ of $\succeq_i$. If u_i is differentiable, then A6 implies that $\delta u_i'(x_i) < u_i'(v_i(x_i, 1))$ whenever $v_i(x_i, 1) > 0$. This condition is weaker than concavity of u_i, which implies $u_i'(x_i) < u_i'(v_i(x_i, 1))$.

This completes our specification of the players' preferences. Since there is no uncertainty explicit in the structure of a bargaining game of alternating offers, and since we restrict attention to situations in which neither player uses a random device to make his choice, there is no need to make assumptions about the players' preferences over uncertain outcomes.

3.3.2 The Intersection of the Graphs of $v_1(\cdot,1)$ and $v_2(\cdot,1)$

In our subsequent analysis the intersection of the graphs of $v_1(\cdot,1)$ and $v_2(\cdot,1)$ has special significance. We now show that this intersection is unique: i.e. there is only one pair $(x,y) \in X \times X$ such that $y_1 = v_1(x_1,1)$ and $x_2 = v_2(y_2,1)$. This uniqueness result is clear from Figure 3.2. Precisely, we have the following.

Lemma 3.2 *If the preference ordering $\succsim_i$ of each Player i satisfies* A2 *through* A6*, then there exists a unique pair $(x^*,y^*) \in X \times X$ such that $y_1^* = v_1(x_1^*,1)$ and $x_2^* = v_2(y_2^*,1)$.*

Proof. For every $x \in X$ let $\psi(x)$ be the agreement for which $\psi_1(x) = v_1(x_1,1)$, and define $H\colon X \to \mathcal{R}$ by $H(x) = x_2 - v_2(\psi_2(x),1)$. The pair of agreements x and $y = \psi(x)$ satisfies also $x_2 = v_2(y_2,1)$ if and only if $H(x) = 0$. We have $H(0,1) \geq 0$ and $H(1,0) \leq 0$, and H is continuous. Hence (by the Intermediate Value Theorem), the function H has a zero. Further, we have

$$H(x) = [v_1(x_1,1) - x_1] + [1 - v_1(x_1,1) - v_2(1 - v_1(x_1,1),1)].$$

Since $v_1(x_1,1)$ is nondecreasing in x_1, both terms are decreasing in x_1 by A6. Thus H has a unique zero. □

The unique pair (x^*,y^*) in the intersection of the graphs is shown in Figure 3.2. Note that this intersection is below the main diagonal, so that $x_1^* > y_1^*$ (and $x_2^* < y_2^*$).

3.3.3 Examples

In subsequent chapters we frequently work with the utility function U_i defined by $U_i(x_i,t) = \delta_i^t x_i$ for every $(x,t) \in X \times T$, and $U_i(D) = 0$, where $0 < \delta_i < 1$. The preferences that this function represents satisfy A1 through A6. We refer to δ_i as the *discount factor* of Player i, and to the preferences as *time preferences with a constant discount rate.*[5] We have $v_i(x_i,t) = \delta_i^t x_i$ in this case, as illustrated in Figure 3.3a.

The utility function defined by $U_i(x_i,t) = x_i - c_i t$ and $U_i(D) = -\infty$, where $c_i > 0$, represents preferences for Player i that satisfy A1 through A5, but not A6. We have $v_i(x_i,t) = x_i - c_i t$ if $x_i \geq c_i t$ and $v_i(x_i,t) = 0$ otherwise (see Figure 3.3b). Thus if $x_i \geq c_i$ then $v_i(x_i,1) = x_i - c_i$, so

[5]This is the conventional name for these preferences. However, given that *any* preferences satisfying A2 through A5 can be represented on $X \times T$ by a utility function of the form $\delta_i^t u_i(x_i)$, the distinguishing feature of time preferences with a constant discount rate is not the constancy of the discount rate but the linearity of the function u_i.

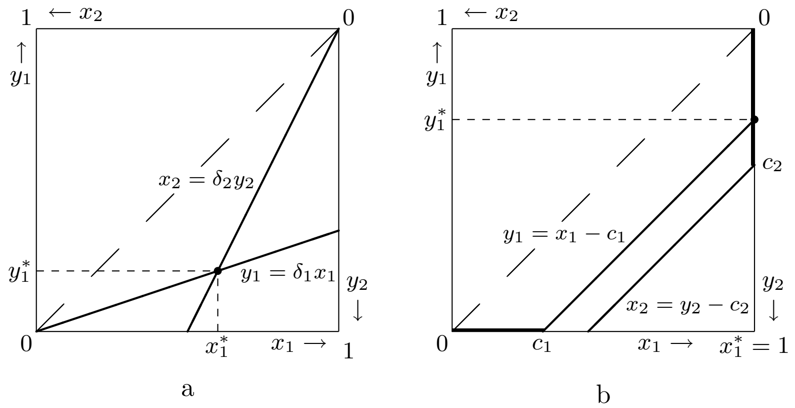


Figure 3.3 Examples of the functions $v_1(\cdot, 1)$ and $v_2(\cdot, 1)$ for (a) time preferences with a constant discount factor and (b) time preferences with a constant cost of delay.

that $x_i - v_i(x_i, 1) = c_i$, which is constant, rather than increasing in x_i. We refer to c_i as the *cost of delay* or *bargaining cost* of Player i, and to the preferences as *time preferences with a constant cost of delay.*

Note that even though preferences with a constant cost of delay violate A6, there is still a unique pair $(x, y) \in X \times X$ such that $y_1 = v_1(x_1, 1)$ and $x_2 = v_2(y_2, 1)$ as long as $c_1 \neq c_2$. Note also that the two families of preferences are qualitatively different. For example, if Player i has time preferences with a constant discount rate then he is indifferent about the timing of an agreement that gives him 0, while if he has time preferences with a constant cost of delay then he prefers to obtain such an agreement as soon as possible. (Since time preferences with a constant cost of delay satisfy A2 through A5, they can be represented on $X \times T$ by a utility function of the form $\delta_i^t u_i(x_i)$ (see the discussion following A5 on p. 34). However, there is no value of δ_i for which u_i is linear.)

3.4 Strategies

A strategy of a player in an extensive game specifies an action at *every* node of the tree at which it is his turn to move.[6] Thus in a bargaining game of alternating offers a strategy of Player 1, for example, begins by specifying (i) the agreement she proposes at $t = 0$, and (ii) for *every* pair consisting

[6]Such a plan of action is sometimes called a *pure* strategy to distinguish it from a plan in which the player uses a random device to choose his action. In this book we allow players to randomize only when we explicitly say so.

of a proposal by Player 1 at $t = 0$ and a counterproposal by Player 2 at $t = 1$, the choice of Y or N at $t = 1$, and, if N is chosen, a further counterproposal for period $t = 2$. The strategy continues by specifying actions at every future period, for *every* possible history of actions up to that point.

More precisely, the players' strategies in a bargaining game of alternating offers are defined as follows. Let X^t be the set of all sequences $(x^0, \ldots, x^{t-1})$ of members of X. A *strategy of Player 1* is a sequence $\sigma = \{\sigma^t\}_{t=0}^{\infty}$ of functions, each of which assigns to each history an action from the relevant set. Thus $\sigma^t\colon X^t \to X$ if t is even, and $\sigma^t\colon X^{t+1} \to \{Y, N\}$ if t is odd: Player 1's strategy prescribes an offer in every even period t for every history of t rejected offers, and a response (accept or reject) in every odd period t for every history consisting of t rejected offers followed by a proposal of Player 2. (The set X^0 consists of the "null" history preceding period 0; formally, it is a singleton, so that σ^0 can be identified with a member of X.) Similarly, a *strategy of Player 2* is a sequence $\tau = \{\tau^t\}_{t=0}^{\infty}$ of functions, with $\tau^t\colon X^{t+1} \to \{Y, N\}$ if t is even, and $\tau^t\colon X^t \to X$ if t is odd: Player 2 accepts or rejects Player 1's offer in every even period, and makes an offer in every odd period.

Note that a strategy specifies actions at every period, for *every* possible history of actions up to that point, *including histories that are precluded by previous actions of Player 1.* Every strategy of Player 1 must, for example, prescribe a choice of Y or N at $t = 1$ in the case that she herself offers $(1/2, 1/2)$ at $t = 0$, and Player 2 rejects this offer and makes a counteroffer, even if the strategy calls for Player 1 to make an offer *different* from $(1/2, 1/2)$ at $t = 0$. Thus Player 1's strategy has to say what she will do at nodes that will never be reached if she follows the prescriptions of her *own* strategy at earlier time periods. At first this may seem strange. In the statement "I will take action x today, and tomorrow I will take action m in the event that I do x today, and n in the event that I do y today", the last clause appears to be superfluous.

If we are interested only in Nash equilibria (see Section 3.6) then there *is* a redundancy in this specification of a strategy. Suppose that the strategy σ' of Player 1 differs from the strategy σ only in the actions it prescribes after histories that are not reached if σ is followed. Then the strategy pairs (σ, τ) and (σ', τ) lead to the same outcome for *every* strategy τ of Player 2. However, if we wish to use the concept of subgame perfect equilibrium (see Section 3.7), then we need a player's strategy to specify his actions after histories that will never occur if he uses that strategy. In order to examine the optimality of Player i's strategy after an arbitrary history—for example, after one in which Player j takes actions inconsistent with his original strategy—we need to invoke Player i's expectation of Player j's

future actions. The components of Player j's strategy that specify his actions after such a history can be interpreted as reflecting j's beliefs about what i expects j to do after this history.

Note that we do not restrict the players' strategies to be "stationary": we allow the players' offers and reactions to offers to depend on events in all previous periods. The assumption of stationarity is sometimes made in models of bargaining, but it is problematic. A stationary strategy is "simple" in the sense that the actions it prescribes in every period do not depend on time, nor on the events in previous periods. However, such a strategy means that Player j expects Player i to adhere to his stationary behavior even if j himself does not. For example, a stationary strategy in which Player 1 always makes the proposal $(1/2, 1/2)$ means that even after Player 1 has made the offer $(3/4, 1/4)$ a thousand times, Player 2 still believes that Player 1 will make the offer $(1/2, 1/2)$ in the next period. If one wishes to assume that the players' strategies are "simple", then it seems that in these circumstances one should assume that Player 2 believes that Player 1 will continue to offer $(3/4, 1/4)$.

3.5 Strategies as Automata

A strategy in a bargaining game of alternating offers can be very complex. The action taken by a player at any point can depend arbitrarily on the entire history of actions up to that point. However, most of the strategies we encounter in the sequel have a relatively simple structure. We now introduce a language that allows us to describe such strategies in a compact and unambiguous way.

The idea is simple. We encode those characteristics of the history that are relevant to a player's choice in a variable called the *state*. A player's action at any point is determined by the state and by the value of some publicly known variables. As play proceeds, the state may change, or it may stay the same; its progression is given by a transition rule. Assigning an action to each of a (typically small) number of states and describing a transition rule is often much simpler than specifying an action after each of the huge number of possible histories.

The publicly known variables include the identity of the player whose turn it is to move and the type of action he has to take (propose an offer or respond to an offer). The progression of these variables is given by the structure of the game. The publicly known variables include also the currently outstanding offer and, in some cases that we consider in later chapters, the most recent rejected offer.

We present our descriptions of strategy profiles in tables, an example of which is Table 3.1. Here there are two states, Q and R. As is our

		State Q	State R
Player 1	proposes	x^Q	x^R
	accepts	$x_1 \geq \alpha$	$x_1 > \beta$
Player 2	proposes	y^Q	y^R
	accepts	$x_1 = 0$	$x_1 < \eta$
	Transitions	Go to R if Player 1 proposes x with $x_1 > \theta$.	Absorbing

Table 3.1 An example of the tables used to describe strategy profiles.

convention, the leftmost column describes the initial state. The first four rows specify the behavior of the players in each state. In state Q, for example, Player 1 proposes the agreement x^Q whenever it is her turn to make an offer and accepts any proposal x for which $x_1 \geq \alpha$ when it is her turn to respond to an offer. The last row indicates the transitions. The entry in this row that lies in the column corresponding to state I $(= Q, R)$ gives the conditions under which there is a transition to a state different from I. The entry "Absorbing" for state R means that there is no transition out of state R: once it is reached, the state remains R forever. As is our convention, every transition occurs immediately after the event that triggers it. (If, for example, in state Q Player 1 proposes x with $x_1 > x_1^Q$, then the state changes to R *before* Player 2 responds.) Note that the same set of states and same transition rule are used to describe both players' strategies. This feature is common to all the equilibria that we describe in this book.

This way of representing a player's strategy is closely related to the notion of an automaton, as used in the theory of computation (see, for example, Hopcroft and Ullman (1979)). The notion of an automaton has been used also in recent work on repeated games; it provides a natural tool to define measures of the complexity of a strategy. Models have been studied in which the players are concerned about the complexity of their strategies, in addition to their payoffs (see, for example, Rubinstein (1986)). Here we use the notion merely as part of a convenient language to describe strategies.

We end this discussion by addressing a delicate point concerning the relation between an automaton as we have defined it and the notion that is used in the theory of computation. We refer to the latter as a "standard automaton". The two notions are not exactly the same, since in our

description a player's action depends not only on the state but also on the publicly known variables. In order to represent players' strategies as standard automata we need to incorporate the publicly known variables into the definitions of the states. The standard automaton that represents Player 1's strategy in Table 3.1, for example, is the following. The set of states is $\{[S,i]: i = 1,2 \text{ and } S = Q,R\} \cup \{[S,i,x]: x \in X,\ i = 1,2, \text{ and } S = Q,R\} \cup \{[x]: x \in X\}$. (The interpretation is that $[S,i]$ is the state in which Player i makes an offer, $[S,i,x]$ is the state in which Player i responds to the offer x, and $[x]$ is the (terminal) state in which the offer x has been accepted.) The initial state is $[Q,1]$. The action Player 1 takes in state $[S,i]$ is the offer specified in column S of the table if $i = 1$ and is null if $i = 2$; the action she takes in state $[S,i,x]$ is either "accept" or "reject", as determined by x and the rule specified for Player i in column S, if $i = 1$, and is null if $i = 2$; and the action she takes in state $[x]$ is null. The transition rule is as follows. If the state is $[S,i,x]$ and the action Player i takes is "reject", then the new state is $[S,i]$; if the action is "accept", then the new state is $[x]$. If the state is $[S,i]$ and the action is the proposal x, then the new state is $[S',j,x]$, where j is the other player and S' is determined by the transition rule given in column S. Finally, if the state is $[x]$ then it remains $[x]$.

3.6 Nash Equilibrium

The following notion of equilibrium in a game is due to Nash (1950b, 1951). A pair of strategies (σ, τ) is a *Nash equilibrium*[7] if, given τ, no strategy of Player 1 results in an outcome that Player 1 prefers to the outcome generated by (σ, τ), and, given σ, no strategy of Player 2 results in an outcome that Player 2 prefers to the outcome generated by (σ, τ).

Nash equilibrium is a standard solution used in game theory. We shall not discuss in detail the basic issue of how it should be interpreted. We have in mind a situation that is stable, in the sense that all players are optimizing given the equilibrium. We do not view an equilibrium necessarily as the outcome of a self-enforcing agreement, or claim that it is a necessary consequence of the players' acting rationally that the strategy profile be a Nash equilibrium. We view the Nash equilibrium as an appropriate solution in situations in which the players are rational, experienced, and have played the same game, or at least similar games, many times.

In some games there is a unique Nash equilibrium, so that the theory gives a very sharp prediction. Unfortunately, this is not so for a bargain-

[7]The only connection between a Nash *equilibrium* and the Nash *solution* studied in Chapter 2 is John Nash.

		$*$
Player 1	proposes	$\overline{x}$
	accepts	$x_1 \geq \overline{x}_1$
Player 2	proposes	$\overline{x}$
	accepts	$x_1 \leq \overline{x}_1$

Table 3.2 A Nash equilibrium of a bargaining game of alternating offers in which the players' preferences satisfy A1 through A6. The agreement $\overline{x}$ is arbitrary.

ing game of alternating offers in which the players' preferences satisfy A1 through A6. In particular, for *every* agreement $x \in X$, the outcome $(x, 0)$ is generated by a Nash equilibrium of such a game.

To show this, let $\overline{x} \in X$ and consider the pair $(\overline{\sigma}, \overline{\tau})$ of (stationary) strategies in which Player 1 *always* proposes $\overline{x}$ and accepts an offer x if and only if $x_1 \geq \overline{x}_1$, and Player 2 *always* proposes $\overline{x}$ and accepts an offer if and only if $x_2 \geq \overline{x}_2$. Formally, for Player 1 let

$$\overline{\sigma}^t(x^0, \ldots, x^{t-1}) = \overline{x} \text{ for all } (x^0, \ldots, x^{t-1}) \in X^t$$

if t is even, and

$$\overline{\sigma}^t(x^0, \ldots, x^t) = \begin{cases} Y & \text{if } x_1^t \geq \overline{x}_1 \\ N & \text{if } x_1^t < \overline{x}_1 \end{cases}$$

if t is odd. Player 2's strategy $\overline{\tau}$ is defined analogously. A representation of $(\overline{\sigma}, \overline{\tau})$ as a pair of (one-state) automata is given in Table 3.2.

If the players use the pair of strategies $(\overline{\sigma}, \overline{\tau})$, then Player 1 proposes $\overline{x}$ at $t = 0$, which Player 2 immediately accepts, so that the outcome is $(\overline{x}, 0)$. To see that $(\overline{\sigma}, \overline{\tau})$ is a Nash equilibrium, suppose that Player i uses a different strategy. Perpetual disagreement is the worst outcome (by A1), and Player j never makes an offer different from x or accepts an agreement x with $x_j < \overline{x}_j$. Thus the best outcome that Player i can obtain, given Player j's strategy, is $(\overline{x}, 0)$.

The set of outcomes generated by Nash equilibria includes not only every possible agreement in period 0, but also some agreements in period 1 or later. Suppose, for example, that $\hat{\sigma}$ and $\hat{\tau}$ differ from $\overline{\sigma}$ and $\overline{\tau}$ only in period 0, when Player 1 makes the offer $(1, 0)$ (instead of $\overline{x}$), and Player 2 rejects every offer. The strategy pair $(\hat{\sigma}, \hat{\tau})$ yields the agreement $(\overline{x}, 1)$, and is an equilibrium if $(\overline{x}, 1) \succsim_2 ((1, 0), 0)$. Unless Player 2 is so impatient that he prefers to receive 0 today rather than 1 tomorrow, there exist values of $\overline{x}$ that satisfy this condition, so that equilibria exist in which agreement is

reached in period 1. A similar argument shows that, for some preferences, there are Nash equilibria in which agreement is reached in period 2, or later.

In summary, the notion of Nash equilibrium puts few restrictions on the outcome in a bargaining game of alternating offers. For this reason, we turn to a stronger notion of equilibrium.

3.7 Subgame Perfect Equilibrium

We can interpret some of the actions prescribed by the strategies $\overline{\sigma}$ and $\overline{\tau}$ defined above as "incredible threats". The strategy $\overline{\tau}$ calls for Player 2 to reject any offer less favorable to him than $\overline{x}$. However, if Player 2 is actually confronted with such an offer, then, under the assumption that Player 1 will otherwise follow the strategy $\overline{\sigma}$, it may be in Player 2's interest to accept the offer rather than reject it. Suppose that $\mathrm{x}_1 < 1$ and that Player 1 makes an offer x in which $x_1 = \overline{x}_1 + \epsilon$ in period t, where $\epsilon > 0$ is small. If Player 2 accepts this offer he receives $\overline{x}_2 - \epsilon$ in period t, while if he rejects it, then, according to the strategy pair $(\overline{\sigma}, \overline{\tau})$, he offers $\overline{x}$ in period $t+1$, which Player 1 accepts, so that the outcome is $(\overline{x}, t+1)$. Player 2 prefers to receive $\overline{x}_2 - \epsilon$ in period t rather than $\overline{x}_2$ in period $t+1$ if ϵ is small enough, so that his "threat" to reject the offer x is not convincing.

The notion of Nash equilibrium does not rule out the use of "incredible threats", because it evaluates the desirability of a strategy only from the viewpoint of the start of the game. As the actions recommended by a strategy pair are followed, a path through the tree is traced out; only a small subset of all the nodes in the tree are reached along this path. The optimality of actions proposed at unreached nodes is not tested when we ask if a strategy pair is a Nash equilibrium. If the two strategies τ and τ' of Player 2 differ only in the actions they prescribe at nodes that are not reached when Player 1 uses the strategy σ, then (σ, τ) and (σ, τ') yield the same path through the tree; hence Player 2 is indifferent between τ and τ' when Player 1 uses σ. To be specific, consider the strategy τ' of Player 2 that differs from the strategy $\overline{\tau}$ defined in the previous section only in period 0, when Player 2 accepts some offers x in which $x_2 < \overline{x}_2$. When Player 1 uses the strategy $\overline{\sigma}$, the strategies $\overline{\tau}$ and τ' generate precisely the same path through the tree—since the strategy $\overline{\sigma}$ calls for Player 1 to offer precisely $\overline{x}$, not an offer less favorable to Player 2. Thus Player 2 is indifferent between $\overline{\tau}$ and τ' when Player 1 uses $\overline{\sigma}$; when considering whether $(\overline{\sigma}, \overline{\tau})$ is a Nash equilibrium we do not examine the desirability of the action proposed by Player 2 in period 0 in the event that Player 1 makes an offer *different* from $\overline{x}$.

Selten's (1965) notion of subgame perfect equilibrium addresses this problem by requiring that a player's strategy be optimal in the game be-

ginning at *every* node of the tree, whether or not that node is reached if the players adhere to their strategies. In the context of the strategy pair $(\overline{\sigma}, \overline{\tau})$ considered in Section 3.6, we ask the following. Suppose that Player 1 makes an offer x different from $\overline{x}$ in period 0. If she otherwise follows the precepts of $\overline{\sigma}$, is it desirable for Player 2 to adhere to $\overline{\tau}$? Since the answer is no when $x_1 = \overline{x}_1 + \epsilon$ and $\epsilon > 0$ is small, the pair $(\overline{\sigma}, \overline{\tau})$ is not a subgame perfect equilibrium. If some strategy pair (σ, τ) passes this test at every node in the tree, then it *is* a subgame perfect equilibrium.

More precisely, for each node of a bargaining game of alternating offers there is an extensive game that starts at this node, which we call a *subgame*.

Definition 3.3 A strategy pair is a *subgame perfect equilibrium* of a bargaining game of alternating offers if the strategy pair it induces in every subgame is a Nash equilibrium of that subgame.

If we represent strategies as (standard) automata (see Section 3.5), then to establish that a strategy profile is a subgame perfect equilibrium it is sufficient to check that no player, in any state, can increase his payoff by a "one-shot" deviation. More precisely, for every pair of (standard) automata and every state there is an outcome associated with the automata if they start to operate in that state in period 0. Since the players' time preferences are stationary (see A5), each player faces a Markovian decision problem, given the other player's automaton.[8] Any change in his strategy that increases his payoff leads to agreement in a *finite* number of periods (given that his preferences satisfy A1), so that his strategy is optimal if, in every state in which he has to move, his action leads to a state for which the outcome is the one he most prefers, among the outcomes in all the states which can be reached by one of his actions.

3.8 The Main Result

We now show that the notion of subgame perfect equilibrium, in sharp contrast to that of Nash equilibrium, predicts a unique outcome in a bargaining game of alternating offers in which the players' preferences satisfy A1 through A6.

The strategies $\overline{\sigma}$ and $\overline{\tau}$ discussed in the previous section call for both players to propose the same agreement $\overline{x}$ and to accept offers only if they are at least as good as $\overline{x}$. Consider an alternative strategy pair $(\hat{\sigma}, \hat{\tau})$ in which Player 1 always (i.e. regardless of the history) offers $\hat{x}$ and accepts an offer y if and only if $y_1 \geq \hat{y}_1$, and Player 2 always offers $\hat{y}$ and accepts an offer x if and only if $x_2 \geq \hat{x}_2$. Under what conditions on $\hat{x}$ and $\hat{y}$ is

[8] For a definition of a Markovian decision problem see, for example, Derman (1970).

$(\hat{\sigma}, \hat{\tau})$ a subgame perfect equilibrium? In the event Player 2 rejects an offer x in period 0, he offers $\hat{y}$ in period 1, which Player 1 accepts. So in order for his rejection of every offer x with $x_2 < \hat{x}_2$ to be credible, we must have $(\hat{y}, 1) \succsim_2 (x, 0)$ whenever $x_2 < \hat{x}_2$; thus if $\hat{x}_2 > 0$ we need $(\hat{y}, 1) \succsim_2 (\hat{x}, 0)$ by continuity (A4). At the same time we must have $(\hat{x}, 0) \succsim_2 (\hat{y}, 1)$, or Player 2 would have an incentive to reject Player 1's offer of $\hat{x}$ in period 0. We conclude that if the strategy pair $(\hat{\sigma}, \hat{\tau})$ is a subgame perfect equilibrium then either $(\hat{x}, 0) \sim_2 (\hat{y}, 1)$, or $\hat{x} = (1, 0)$ and $(\hat{x}, 0) \succsim_2 (\hat{y}, 1)$; or, stated more compactly, $v_2(\hat{y}_2, 1) = \hat{x}_2$ (see (3.1)). Applying a similar logic to Player 1's rule for accepting offers in period 1, we conclude that we need either $(\hat{y}, 1) \sim_1 (\hat{x}, 2)$, or $\hat{y} = (0, 1)$ and $(\hat{y}, 1) \succsim_1 (\hat{x}, 2)$. By our stationarity assumption (A5), this is equivalent to $v_1(\hat{x}_1, 1) = \hat{y}_1$.

This argument shows that if $(\hat{\sigma}, \hat{\tau})$ is a subgame perfect equilibrium then $(\hat{x}, \hat{y})$ must coincide with the unique solution (x^*, y^*) of the following equations.

$$y_1^* = v_1(x_1^*, 1) \qquad \text{and} \qquad x_2^* = v_2(y_2^*, 1). \tag{3.3}$$

(The uniqueness follows from Lemma 3.2.) Note that if $y_1^* > 0$ and $x_2^* > 0$ then

$$(y^*, 0) \sim_1 (x^*, 1) \qquad \text{and} \qquad (x^*, 0) \sim_2 (y^*, 1). \tag{3.4}$$

Note further that if the players' preferences are such that for each Player i and every outcome (x, t) there is an agreement y such that Player i is indifferent between $(y, 0)$ and (x, t), then in the unique solution (x^*, y^*) of (3.3) we have $y_1^* > 0$ and $x_2^* > 0$, so that (x^*, y^*) satisfies (3.4).

The main result of this chapter is that any bargaining game of alternating offer in which the players' preferences satisfy A1 through A6 has a unique subgame perfect equilibrium, which has the structure of $(\hat{\sigma}, \hat{\tau})$.

Theorem 3.4 *Every bargaining game of alternating offers in which the players' preferences satisfy A1 through A6 has a unique subgame perfect equilibrium (σ^*, τ^*). In this equilibrium Player 1 proposes the agreement x^* defined in (3.3) whenever it is her turn to make an offer, and accepts an offer y of Player 2 if and only if $y_1 \geq y_1^*$; Player 2 always proposes y^*, and accepts only those offers x with $x_2 \geq x_2^*$. The outcome is that Player 1 proposes x^* in period 0, and Player 2 immediately accepts this offer.*

Formally, the subgame perfect equilibrium strategy σ^* of Player 1 described in the theorem is defined by

$$\sigma^{*t}(x^0, \ldots, x^{t-1}) = x^* \text{ for all } (x^0, \ldots, x^{t-1}) \in X^t$$

if t is even, and

$$\sigma^{*t}(x^0, \ldots, x^t) = \begin{cases} Y & \text{if } x_1^t \geq y_1^* \\ N & \text{if } x_1^t < y_1^* \end{cases}$$

		$*$
Player 1	proposes	x^*
	accepts	$x_1 \geq y_1^*$
Player 2	proposes	y^*
	accepts	$x_1 \leq x_1^*$

Table 3.3 The unique subgame perfect equilibrium of a bargaining game of alternating offers in which the players' preferences satisfy A1 through A6. The pair of agreements (x^*, y^*) is the unique solution of (3.3).

if t is odd. The strategy τ^* of Player 2 has the same structure; the roles of x^* and y^* are reversed, the words "odd" and "even" are interchanged, and each subscript 1 is replaced by 2. Table 3.3 describes the strategies σ^* and τ^* as automata.

Note the we have not *assumed* that the strategies are stationary; we have allowed actions in any period to depend on the entire history of the game. The theorem establishes that the only subgame perfect equilibrium strategies take this form.

Proof of Theorem 3.4. First we argue that the strategy pair (σ^*, τ^*) is a subgame perfect equilibrium. We need to show that (σ^*, τ^*) induces a Nash equilibrium in every subgame. Consider a subgame starting with an offer by Player 1 in period t^*. Given that Player 2 uses the strategy τ^*, any strategy of Player 1 that proposes x^* in period t^* leads to the outcome (x^*, t^*); any other strategy of Player 1 generates either (x, t) where $x_1 \leq x_1^*$ and $t \geq t^*$, or (y^*, t) where $t \geq t^* + 1$, or D. Since $x_1^* > y_1^*$, it follows from A1, A2, and A3 that the best of these outcomes for Player 1 is (x^*, t^*), so that σ^* is a best response to τ^* in the subgame. Given that Player 1 uses the strategy σ^*, any strategy of Player 2 that accepts x^* in period t^* leads to the outcome (x^*, t^*); any other strategy of Player 2 generates either (x^*, t) for $t > t^*$, or (y, t) where $y_2 \leq y_2^*$ and $t \geq t^* + 1$, or D. By A1, A2, and A3 the best of these outcomes for Player 2 is either (x^*, t^*) or $(y^*, t^* + 1)$. Now, by definition we have $x_2^* = v_2(y_2^*, 1)$, so that $(x^*, 0) \succsim_2 (y^*, 1)$ (see (3.2)), and hence by (A5) (stationarity), $(x^*, t^*) \succsim_2 (y^*, t^* + 1)$. Thus τ^* is a best response for Player 2 to σ^* in the subgame. Similar arguments apply to subgames starting with an offer by Player 2 and to subgames starting with a response by either player.

We now turn to the more difficult part of the argument, which shows that (σ^*, τ^*) is the only subgame perfect equilibrium.

For $i = 1, 2$, all subgames that begin with an offer by Player i are isomorphic (by the stationarity assumption A5); let G_i be such a subgame. The existence of the SPE above allows us to define

$$M_i = \sup\{v_i(x_i, t)\colon \text{ there is an SPE of } G_i \text{ with outcome } (x, t)\},$$

where SPE means "subgame perfect equilibrium"; let m_i be the corresponding infimum. Note that M_1 and m_1 are defined on a subgame beginning with an offer by Player 1, while M_2 and m_2 are defined on a subgame beginning with an offer by Player 2. We shall show that

$$M_1 = m_1 = x_1^* \qquad \text{and} \qquad M_2 = m_2 = y_2^*, \tag{3.5}$$

so that the present value for Player 1 of every SPE outcome of G_1 is x_1^*, and the present value for Player 2 of every SPE outcome of G_2 is y_2^*. By the following argument, this suffices to prove the theorem.

We need to show that it follows from (3.5) that every SPE of G_1 is (σ^*, τ^*). First we argue that in any SPE the first offer is accepted. Suppose to the contrary that there is an SPE in which Player 1's first offer x is rejected. After the rejection, the players must follow an SPE of G_2. By (3.5) the present value to Player 2 of such an SPE is y_2^*, so that the present value to Player 1 is no more than y_1^*. Since $v_1(y_1^*, 1) \leq y_1^* < x_1^*$, the present value of the SPE to Player 1 is less than x_1^*, contradicting (3.5). Thus in every SPE of G_1 the first offer is accepted. A similar argument applies to G_2. It follows that in any SPE of G_1, Player 1 always proposes x^*, which Player 2 accepts, and Player 2 always proposes y^*, which Player 1 accepts. Also, by (3.3), Player 1 rejects any offer y in which $y_1 < y_1^*$ and accepts any offer y in which $y_1 > y_1^*$; analogously for Player 2.

It remains to establish (3.5). We do so in two steps.

Step 1. $m_2 \geq 1 - v_1(M_1, 1)$.

Proof. Suppose that in the first period of G_2 Player 2 proposes z with $z_1 > v_1(M_1, 1)$. If Player 1 accepts z then the outcome is $(z, 0)$. If she rejects z, then the outcome has present value at most $v_1(M_1, 1)$ to her. Thus in any SPE she accepts any such proposal z, and hence $m_2 \geq 1 - v_1(M_1, 1)$.

Step 2. $M_1 \leq 1 - v_2(m_2, 1)$.

Proof. If, in the first period of G_1, Player 2 rejects the offer of Player 1, then he can obtain at least m_2 with one period of delay. Hence in any SPE Player 2 rejects any offer x for which $x_2 < v_2(m_2, 1)$. Thus the most that Player 1 can obtain if agreement is reached in the first period is $1 - v_2(m_2, 1)$. Since the outcome in any SPE in which agreement is delayed has present value to Player 1 no greater than $v_1(1 - m_2, 1) \leq 1 - m_2 \leq 1 - v_2(m_2, 1)$, the result follows.

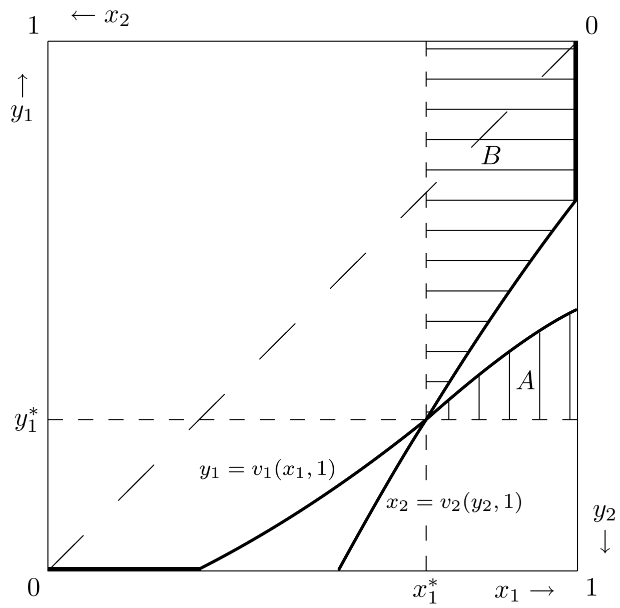


Figure 3.4 An illustration of the last part of the proof of Theorem 3.4. It follows from Step 1 and the fact that $m_2 \le y_2^*$ that the pair $(M_1, 1 - m_2)$ lies in the region labeled A; it follows from Step 2 and the fact that $M_1 \ge x_1^*$ that this pair lies in the region labeled B.

Step 1 establishes that in Figure 3.4 the pair $(M_1, 1 - m_2)$ (relative to the origin at the bottom left) lies below the line $y_1 = v_1(x_1, 1)$. Similarly, Step 2 establishes that $(M_1, 1-m_2)$ lies to the left of the line $x_2 = v_2(y_2, 1)$. Since we showed in the first part of the proof that (σ^*, τ^*) is an SPE of G_1, we know that $M_1 \ge x_1^*$; the same argument shows that there is an SPE of G_2 in which the outcome is $(y^*, 0)$, so that $m_2 \le y_2^*$, and hence $1 - m_2 \ge y_1^*$. Combining these facts we conclude from Figure 3.4 that $M_1 = x_1^*$ and $m_2 = y_2^*$.

The same arguments, with the roles of the players reversed, show that $m_1 = x_1^*$ and $M_2 = y_2^*$. This establishes (3.5), completing the proof. □

The proof relies heavily on the fact that there is a unique solution to (3.3) but does not otherwise use the condition of increasing loss to delay (A6) which we imposed on preferences. Thus any other condition that guarantees a unique solution to (3.3) can be used instead of A6.

3.9 Examples

3.9.1 Constant Discount Rates

Suppose that the players have time preferences with constant discount rates (i.e. Player i's preferences over outcomes (x, t) are represented by the utility function $\delta_i^t x_i$, where $\delta_i \in (0, 1)$ (see Section 3.3.3)). Then (3.3) implies that $y_1^* = \delta_1 x_1^*$ and $x_2^* = \delta_2 y_2^*$, so that

$$x^* = \left(\frac{1-\delta_2}{1-\delta_1\delta_2}, \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2}\right) \quad \text{and} \quad y^* = \left(\frac{\delta_1(1-\delta_2)}{1-\delta_1\delta_2}, \frac{1-\delta_1}{1-\delta_1\delta_2}\right). \quad (3.6)$$

Thus if $\delta_1 = \delta_2 = \delta$ (the discount factors are equal), then $x^* = (1/(1+\delta), \delta/(1+\delta))$.

Notice that as δ_1 approaches 1, the agreement x^* approaches $(1, 0)$: as Player 1 becomes more patient, her share increases, and, in the limit, she receives all the pie. Similarly, as Player 2 becomes more patient, Player 1's share of the pie approaches zero. The cases in which δ_1 or δ_2 are actually equal to 1 are excluded by assumption A3. Nevertheless, if only one of the δ_i's is equal to one then the proof that there is a unique subgame perfect equilibrium payoff vector is still valid, although in this case there is a multiplicity of subgame perfect equilibria. For example, if $\delta_1 = 1$ and $\delta_2 < 1$, then the unique subgame perfect equilibrium payoff vector is $(1, 0)$, but in addition to the equilibrium described in Theorem 3.4 there is one in which Player 2 rejects the offer $(1, 0)$ in period 0 and proposes $(1, 0)$ in period 1, which Player 1 accepts.

3.9.2 Constant Costs of Delay

Preferences that display constant costs of delay are represented by the utility function $x_i - c_i t$, where $c_i > 0$. As remarked in Section 3.3.3, these preferences do not satisfy assumption A6. Nevertheless, as long as $c_1 \neq c_2$ there is a unique pair (x^*, y^*) that satisfies (3.3): $x^* = (1, 0)$ and $y^* = (1 - c_1, c_1)$ if $c_1 < c_2$ (see Figure 3.3b), and $x^* = (c_2, 1 - c_2)$ and $y^* = (0, 1)$ if $c_1 > c_2$. Thus, because of the remark following the proof, Theorem 3.4 still applies: there is a unique subgame perfect equilibrium, in which the players immediately reach the agreement $x^* = (1, 0)$ if $c_1 < c_2$, and $x^* = (c_2, 1 - c_2)$ if $c_1 > c_2$. The prediction here is quite extreme—Player 1 gets all the pie if her delay cost is smaller than that of Player 2, while Player 2 gets $1 - c_2$ if his delay cost is smaller. When the delay costs are the same and less than 1, there is no longer a unique solution to (3.3); in this case there are multiple subgame perfect equilibria if the delay cost is small enough, and in some of these equilibria agreement is not reached in period 0 (see Rubinstein (1982, pp. 107–108)).

3.10 Properties of the Subgame Perfect Equilibrium

3.10.1 Delay

The structure of a bargaining game of alternating offers allows negotiation to continue indefinitely. Nevertheless, in the unique subgame perfect equilibrium it terminates immediately; from an economic point of view, the bargaining process is efficient (no resources are lost in delay). To which features of the model can we attribute this result? We saw that in a Nash equilibrium of the game, delay is possible. Thus the notion of subgame perfection plays a role in the result. Yet perfection alone does not rule out delay—our assumptions on preferences are also important.

To see this, notice that the proof that agreement is reached immediately if the game has a unique subgame perfect equilibrium payoff vector relies only on assumptions A1, A2, and A3. In other words, if the players' preferences satisfy these three assumptions and there is a unique subgame perfect equilibrium then there is no delay. Thus the presence of delay is closely related to the existence of multiple equilibria, which arises, for example, if both players' time preferences have the same constant cost of delay (see Section 3.9.2). It is convenient to demonstrate this point by considering another case in which there is a multiplicity of equilibria.

Suppose that there are just three divisions of the pie available: $X = \{a, b, c\}$. Assume that $a_1 > b_1 > c_1$, and that the players' preferences satisfy A1, A2, A3, and A5. Further assume that if a player prefers (x, t) to (y, t), then he also prefers $(x, t+1)$ to (y, t) (so that $(a, 1) \succ_1 (b, 0)$, $(b, 1) \succ_1 (c, 0)$, $(b, 1) \succ_2 (a, 0)$, and $(c, 1) \succ_2 (b, 0)$). Then for each x$\in X$, the pair of strategies in which each player always insists on $\overline{x}$ (i.e. Player i always offers $\overline{x}$ and accepts an offer x if and only if $x_i \geq \overline{x}_i$) is a subgame perfect equilibrium.

We now construct a subgame perfect equilibrium in which agreement is reached in period 1. In period 0, Player 1 proposes a. Player 2 rejects an offer of a or b, and accepts c. If Player 1 offers a in period 0 and this is rejected, then from period 1 on the subgame perfect equilibrium strategy pair in which each player insists on b (as described above) is played. If Player 1 offers b or c in period 0 and this is rejected, then from period 1 on the subgame perfect equilibrium strategy pair in which each player insists on c is played. These strategies are described in Table 3.4 as automata. There are three states, A, B, and C; as is our convention, the leftmost state (A) is the initial state. (Since it is not possible to reach a situation in which the state is A and either Player 1 has to respond to an offer or Player 2 has to make an offer, the corresponding boxes in the table are blank.)

		A	B	C
1	proposes	a	b	c
	accepts		a and b	a, b, and c
2	proposes		b	c
	accepts	c	b and c	c
Transitions		Go to B if Player 2 rejects a.	Absorbing	Absorbing
		Go to C if Player 2 rejects b or c.		

Table 3.4 A subgame perfect equilibrium of a bargaining game of alternating offers in which there are only three divisions of the pie available. It is not possible to reach a situation in which the state is A and either Player 1 has to respond to an offer or Player 2 has to make an offer, so that the corresponding entries are blank.

The outcome of this strategy profile is that Player 1 offers a in period 0, and Player 2 rejects this offer and proposes b in period 1, which Player 1 accepts. To check that the strategies constitute a subgame perfect equilibrium, notice that if Player 1 offers b rather than a, then the outcome is $(c, 1)$, which is worse for her than $(b, 1)$. If she offers c then the outcome is $(c, 0)$, which is also worse for her than $(b, 1)$.

A final ingredient of the model that appears to contribute to the result that an agreement is reached without delay is the basic assumption that each player is completely informed about the preferences of his opponent. Intuition suggests that if a player is uncertain about his opponent's characteristics then negotiation could be lengthy: a player might make an offer that is accepted by some kinds of opponent and rejected by others. We return to this issue in Chapter 5.

3.10.2 Patience

The equilibrium outcome depends on the character of the players' preferences. One characteristic that we can isolate is the degree of *patience*. Define the preferences $\succsim_1'$ to be *less patient than* $\succsim_1$ if $v_1'(x_1, 1) \le v_1(x_1, 1)$ for all $x \in X$, and $v_1'(x_1, 1) < v_1(x_1, 1)$ for some $x \in X$. It is immediate from a diagram like that in Figure 3.2 that the value of x_1^* that solves (3.3) for the preferences $\succsim_1'$ is no larger than the value that solves (3.3) for the preferences $\succsim_1$, and may be smaller. Thus the model pre-

dicts that when a player becomes less patient, his negotiated share of the pie decreases.

If the players have time preferences with constant discount rates, then being less patient means having a smaller value of δ_i. In this case we can read off the result from (3.6): if δ_1 decreases then x_1^* decreases, while if δ_2 decreases then x_1^* increases.

3.10.3 Symmetry

The structure of a bargaining game of alternating offers is asymmetric in one respect: one of the bargainers is the first to make an offer. If the player who starts the bargaining has the preferences $\succsim_2$ while the player who is the first to respond has the preferences $\succsim_1$, then Theorem 3.4 implies that in the only subgame perfect equilibrium the players reach the agreement y^* (see (3.3)) in period 0. Since $x_1^* > y_1^*$, being the first to make an offer gives a player an advantage. If the players' attitudes to time are the same then we can be more specific. In this case $v_1 = v_2$, so that in the solution to (3.3) we have $x_1^* = y_2^* = 1 - y_1^*$. Given that $x_1^* > y_1^*$ we have $x_1^* > 1/2$ and $y_1^* < 1/2$: the first to move obtains more than half of the pie.

In a game in which one player makes all the offers, there is a unique subgame perfect equilibrium, in which that player obtains all the pie (regardless of the players' preferences). The fact that the player who makes the first offer has an advantage when the players alternate offers is a residue of the extreme asymmetry when one player alone makes all the offers.

The asymmetry in the structure of a bargaining game of alternating offers is artificial. One way of diminishing its effect is to reduce the amount of time that elapses between periods. In Section 4.4 we consider the effect of doing so for a wide class of preferences. Here we simply note what happens when the players have time preferences with constant discount rates. In this case we can simulate the effect of shrinking the length of the period by considering a sequence of games indexed by Δ in which Player i's utility for the agreement x reached after a delay of t periods is $\delta_i^{\Delta t} x_i$. Let $x^*(\Delta)$ be the agreement reached (in period 0) in the unique subgame perfect equilibrium of the game indexed by Δ in which Player 1 is the first to make an offer. Let $y^*(\Delta)$ be the agreement reached in this game when Player 2 is the first to make an offer. It follows from the calculations in Section 3.9.1 that $x_1^*(\Delta) = (1 - \delta_2^\Delta)/(1 - \delta_1^\Delta \delta_2^\Delta)$ and $y_2^*(\Delta) = (1 - \delta_1^\Delta)/(1 - \delta_1^\Delta \delta_2^\Delta)$. Using l'Hôpital's rule we find that

$$\lim_{\Delta \to 0} x^*(\Delta) = \lim_{\Delta \to 0} y^*(\Delta) = \left(\frac{\log \delta_2}{\log \delta_1 + \log \delta_2}, \frac{\log \delta_1}{\log \delta_1 + \log \delta_2} \right).$$

Thus the limit, as the length of the period shrinks to 0, of the amount received by a player is the same regardless of which player makes the first offer.

As an alternative to shrinking the length of the period, we can modify the game to make its structure symmetric. One way of doing so is to consider a game in which at the beginning of each period each player is chosen with probability 1/2 (independently across periods) to be the one to make an offer. Since this introduces uncertainty into the structure of the game, we need to make assumptions about the players' preferences among lotteries over outcomes. If we make the assumptions of von Neumann and Morgenstern then we can show that this game has a unique subgame perfect equilibrium. In this equilibrium, Player 1 always offers $\tilde{x}$ and Player 2 always offers $\tilde{y}$, where $(\tilde{x}, \tilde{y})$ is such that Player 1 is indifferent between $(\tilde{y}, 0)$ and the lottery that yields $(\tilde{x}, 1)$ and $(\tilde{y}, 1)$ with equal probabilities, and Player 2 is indifferent between $(\tilde{x}, 0)$ and the same lottery. (We omit the details.)

3.10.4 Stationarity of Preferences

Theorem 3.4 continues to hold if we weaken assumption A5 and require only that Player 1's preference between the outcomes (x, t) and $(y, t+1)$ when t is *odd* is independent of t, and Player 2's preference between (x, t) and $(y, t+1)$ when t is *even* is independent of t. The reason is that in addition to A1, A2, and A3, the only property of preferences that we have used concerns the players' preference between accepting an offer and rejecting it and thus moving the play into a subgame starting in the next period. Thus Player 1's preference between (x, t) and $(y, t+1)$ when t is even, and Player 2's preference between these outcomes when t is odd, are irrelevant. As long as the preferences continue to satisfy A1, A2, and A3, there is a unique subgame perfect equilibrium, which is characterized by a suitably modified version of (3.3):

$$v_1(y_1^*, 1) = v_1(x_1^*, 2) \qquad \text{and} \qquad x_2^* = v_2(y_2^*, 1). \tag{3.7}$$

To illustrate this point, consider the case in which each period corresponds to an interval of real time. Suppose that Player i's preferences over pairs (x, θ), where x is an agreement and θ is the real time at which the agreement is reached, are represented by the utility function $\delta^{\theta} x_i$. Assume that the time it takes Player i to make a new proposal after he rejects one is Δ_i. Then the unique subgame perfect equilibrium of this game is the same as the unique subgame perfect equilibrium of the game in which each period has length 1 and the players have constant discount factors δ^{Δ_i}. The more quickly Player i can make a counterof-

fer after rejecting an offer of Player j, the larger is δ^{Δ_i}, and hence the larger is x_1^* and the smaller is y_1^*. In the limit, when Player 1 can respond instantly ($\Delta_1 = 0$), but Player 2 cannot, Player 1 obtains all the pie ($x^* = y^* = (1, 0)$). In Section 4.4.4 we further study the case of asymmetric response times.

3.11 Finite versus Infinite Horizons

Our choice of an infinite horizon for the bargaining game raises an important modeling issue. At first glance the assumption of an infinite horizon is not realistic: every individual's life is finite. As an alternative, we can construct a model in which the horizon is either some fixed finite number or a random variable with a finite support.

A bargaining game of alternating offers with a finite horizon has a unique subgame perfect equilibrium (under the assumptions on preferences made in Section 3.3), which can be calculated by backwards induction. As the horizon increases, the agreement reached in this equilibrium converges to the agreement reached in the unique subgame perfect equilibrium of the model with an infinite horizon. (Binmore (1987b) uses this fact to provide an alternative proof of Theorem 3.4.) Thus the infinite horizon model of this chapter predicts an outcome very similar to that predicted by a model with a very long finite horizon.

Despite the similarity in the predictions of the models, we do not regard the differences between the models as insignificant. The model with an infinite horizon fits a situation in which the players perceive that, after any rejection of an offer, there is room for a counterproposal. Such a perception ignores the fact that the death of one of the players or the end of the world may preclude any counterproposal. The model with a finite horizon fits a situation in which the final stage of the game is perceived clearly by the players, who fully take it into account when formulating their strategies.

The significant difference between the two models lies not in the realism of the horizons they posit but in the strategic reasoning of the players. In many contexts a model in which the horizon is infinite better captures this reasoning process. In such cases, a convergence theorem for games with finite horizons may be useful as a technical device, even if the finite games themselves are of limited intrinsic interest.

3.12 Models in Which Players Have Outside Options

Here we analyze two modifications of the structure of a bargaining game of alternating offers in which one of the players has the option of leaving

his current partner, in which case the game ends. In both cases we restrict attention to the case in which the players have time preferences with the same constant discount factor $\delta < 1$.

We consider two games, in each of which Player 2 has the option of terminating the negotiation; in this event an outcome that is worth b to him (and 0 to Player 1) occurs. The games differ in the times at which it is possible for Player 2 to quit. If he can quit only after he has rejected an offer, then the game has a unique subgame perfect equilibrium. If he can quit either only after Player 1 rejects his offer or after any rejection, then, for some values of the outside option, the game has multiple subgame perfect equilibria. In either case, if b is small relative to the payoff of Player 2 in the unique subgame perfect equilibrium of the game in which there is no outside option, then this outside option has no effect on the outcome of the game. This result is striking. An intuition for it is that opting out is not a credible threat for Player 2: he can achieve no more outside the relationship than he can within it. If b is large, then in the first model there is a unique subgame perfect equilibrium in which the players obtain the payoffs $(1 - b, b)$, while in the second model there is a range of subgame perfect equilibrium payoffs.

3.12.1 A Model in Which Player 2 Can Opt Out Only When Responding to an Offer

We study a modification of the model of alternating offers in which Player 2, and only Player 2, can unilaterally quit the negotiation. If this event (the "outside option") occurs in period t then the players obtain the utility pair $(0, \delta^t b)$, where $b < 1$. If $b > 0$ then Player 2 seems to have an advantage over Player 1. He has a valuable alternative to reaching an agreement with Player 1, while Player 1 has no choice but to bargain with Player 2.

When can Player 2 opt out? It turns out that this question is important. In this section we assume that Player 2 can opt out *only* when responding to an offer from Player 1. The structure of negotiation is thus the following. First Player 1 proposes a division x of the pie. Player 2 may accept this proposal, reject it and opt out, or reject it and continue bargaining. In the first two cases the negotiation ends; in the first case the payoff vector is x, and in the second case it is $(0, b)$. If Player 2 rejects the offer and continues bargaining, play passes into the next period, when it is Player 2's turn to make an offer, which Player 1 may accept or reject. In the event of rejection, another period passes, and once again it is Player 1's turn to make an offer. The first two periods of the resulting game are shown in Figure 3.5. The result we obtain is the following.

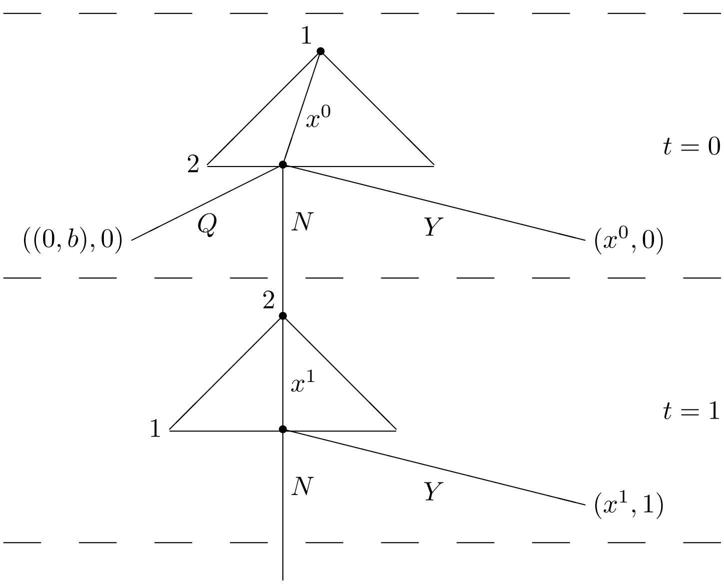


Figure 3.5 The first two periods of a bargaining game in which Player 2 can opt out only when responding to an offer. The branch labelled x^0 represents a "typical" offer of Player 1 out of the continuum available in period 0; similarly, the branch labeled x^1 is a "typical" offer of Player 2 in period 1. In period 0, Player 2 can reject the offer and opt out (Q), reject the offer and continue bargaining (N), or accept the offer (Y).

Proposition 3.5 *Consider the bargaining game described above, in which Player 2 can opt out only when responding to an offer, as in Figure 3.5. Assume that the players have time preferences with the same constant discount factor $\delta < 1$, and that their payoffs in the event that Player 2 opts out in period t are $(0, \delta^t b)$, where $b < 1$.*

1. *If $b < \delta/(1+\delta)$ then the game has a unique subgame perfect equilibrium, which coincides with the subgame perfect equilibrium of the game in which Player 2 has no outside option. That is, Player 1 always proposes the agreement $(1/(1+\delta), \delta/(1+\delta))$ and accepts any proposal y in which $y_1 \geq \delta/(1+\delta)$, and Player 2 always proposes the agreement $(\delta/(1+\delta), 1/(1+\delta))$, accepts any proposal x in which $x_2 \geq \delta/(1+\delta)$, and never opts out. The outcome is that agreement is reached immediately on $(1/(1+\delta), \delta/(1+\delta))$.*

2. *If $b > \delta/(1+\delta)$ then the game has a unique subgame perfect equilibrium, in which Player 1 always proposes $(1-b, b)$ and accepts any*

proposal y in which $y_1 \geq \delta(1-b)$, and Player 2 always proposes $(\delta(1-b), 1-\delta(1-b))$, accepts any proposal x in which $x_2 \geq b$, and opts out if $x_2 < b$. The outcome is that agreement is reached immediately on the division $(1-b, b)$.

3. *If $b = \delta/(1+\delta)$ then in every subgame perfect equilibrium the outcome is an immediate agreement on $(1-b, b)$.*

Proof. Throughout this proof we write SPE for "subgame perfect equilibrium". First note that if $\delta/(1+\delta) \geq b$ then the SPE of the bargaining game of alternating offers given in Theorem 3.4 is an SPE of the game here. (Given the equilibrium strategies, Player 2 can never improve his position by opting out.)

If $\delta/(1+\delta) \leq b$ then the argument that the pair of strategies given in Part 2 of the proposition is an SPE is straightforward. For example, to check that it is optimal for Player 2 to opt out when responding to an offer x with $x_2 < b$ in period t, consider the payoffs from his three possible actions. If he opts out, he obtains b; if he accepts the offer, he obtains $x_2 < b$. If he rejects the offer and continues bargaining then the best payoff he can obtain in period $t+1$ is $1-\delta(1-b)$, and the payoff he can obtain in period $t+2$ is b. Because of the stationarity of Player 1's strategy, Player 2 is worse off if he waits beyond period $t+2$. Now, we have $\delta^2 b \leq \delta[1-\delta(1-b)] \leq b$ (the second inequality since $\delta/(1+\delta) \leq b$). Thus Player 2's optimal action is to opt out if Player 1 proposes an agreement x in which $x_2 < b$.

Let M_1 and M_2 be the suprema of Player 1's and Player 2's payoffs over SPEs of the subgames in which Players 1 and 2, respectively, make the first offer. Similarly, let m_1 and m_2 be the infima of these payoffs. We proceed in a number of steps.

Step 1. $m_2 \geq 1 - \delta M_1$.

The proof is the same as that of Step 1 in the proof of Theorem 3.4.

Step 2. $M_1 \leq 1 - \max\{b, \delta m_2\}$.

Proof. Since Player 2 obtains the utility b by opting out, we must have $M_1 \leq 1 - b$. The fact that $M_1 \leq 1 - \delta m_2$ follows from the same argument as for Step 2 in the proof of Theorem 3.4.

Step 3. $m_1 \geq 1 - \max\{b, \delta M_2\}$ and $M_2 \leq 1 - \delta m_1$.

The proof is analogous to those for Steps 1 and 2.

Step 4. If $\delta/(1+\delta) \geq b$ then $m_i \leq 1/(1+\delta) \leq M_i$ for $i = 1, 2$.

Proof. These inequalities follow from the fact that in the SPE described in the proposition Player 1 obtains the utility $1/(1+\delta)$ in any subgame

in which she makes the first offer, and Player 2 obtains the same utility in any subgame in which he makes the first offer.

Step 5. If $\delta/(1+\delta) \geq b$ then $M_1 = m_1 = 1/(1+\delta)$ and $M_2 = m_2 = 1/(1+\delta)$.

Proof. By Step 2 we have $1 - M_1 \geq \delta m_2$, and by Step 1 we have $m_2 \geq 1 - \delta M_1$, so that $1 - M_1 \geq \delta - \delta^2 M_1$, and hence $M_1 \leq 1/(1+\delta)$. Hence $M_1 = 1/(1+\delta)$ by Step 4.

Now, by Step 1 we have $m_2 \geq 1 - \delta M_1 = 1/(1+\delta)$. Hence $m_2 = 1/(1+\delta)$ by Step 4.

Again using Step 4 we have $\delta M_2 \geq \delta/(1+\delta) \geq b$, and hence by Step 3 we have $m_1 \geq 1 - \delta M_2 \geq 1 - \delta(1 - \delta m_1)$. Thus $m_1 \geq 1/(1+\delta)$. Hence $m_1 = 1/(1+\delta)$ by Step 4.

Finally, by Step 3 we have $M_2 \leq 1 - \delta m_1 = 1/(1+\delta)$, so that $M_2 = 1/(1+\delta)$ by Step 4.

Step 6. If $b \geq \delta/(1+\delta)$ then $m_1 \leq 1-b \leq M_1$ and $m_2 \leq 1-\delta(1-b) \leq M_2$.

Proof. These inequalities follow from the SPE described in the proposition (as in Step 4).

Step 7. If $b \geq \delta/(1+\delta)$ then $M_1 = m_1 = 1-b$ and $M_2 = m_2 = 1-\delta(1-b)$.

Proof. By Step 2 we have $M_1 \leq 1 - b$, so that $M_1 = 1 - b$ by Step 6. By Step 1 we have $m_2 \geq 1 - \delta M_1 = 1 - \delta(1-b)$, so that $m_2 = 1 - \delta(1-b)$ by Step 6.

Now we show that $\delta M_2 \leq b$. If $\delta M_2 > b$ then by Step 3 we have $M_2 \leq 1 - \delta m_1 \leq 1 - \delta(1 - \delta M_2)$, so that $M_2 \leq 1/(1+\delta)$. Hence $b < \delta M_2 \leq \delta/(1+\delta)$, contradicting our assumption that $b \geq \delta/(1+\delta)$.

Given that $\delta M_2 \leq b$ we have $m_1 \geq 1 - b$ by Step 3, so that $m_1 = 1 - b$ by Step 6. Further, $M_2 \leq 1 - \delta m_1 = 1 - \delta(1-b)$ by Step 3, so that $M_2 = 1 - \delta(1-b)$ by Step 6.

Thus in each case the SPE outcome is unique. The argument that the SPE strategies are unique if $b \neq \delta/(1+\delta)$ is the same as in the proof of Theorem 3.4. If $b = \delta/(1+\delta)$ then there is more than one SPE; in some SPEs, Player 2 opts out when facing an offer that gives him less than b, while in others he continues bargaining in this case. □

3.12.2 *A Model in Which Player 2 Can Opt Out Only After Player 1 Rejects an Offer*

Here we study another modification of the bargaining game of alternating offers. In contrast to the previous section, we assume that Player 2 may opt

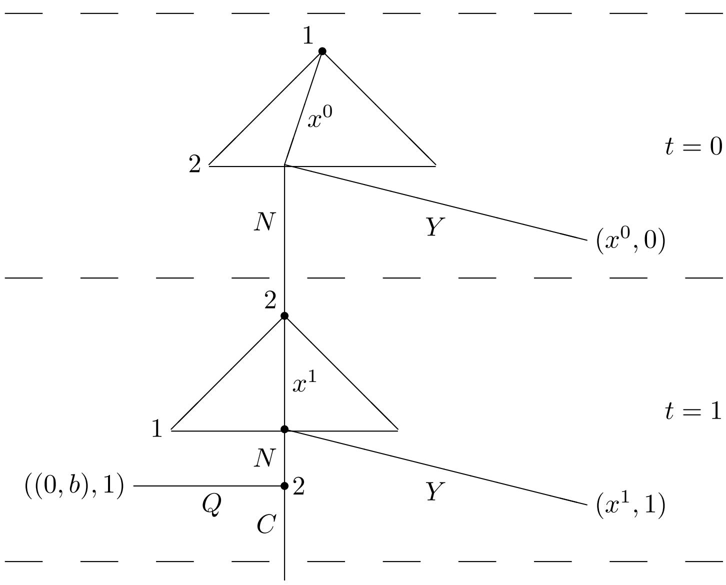


Figure 3.6 The first two periods of a bargaining game in which Player 2 can opt out only after Player 1 rejects an offer. The branch labelled x^0 represents a "typical" offer of Player 1 out of the continuum available in period 0; similarly, the branch labeled x^1 is a "typical" offer of Player 2 in period 1. In period 0, Player 2 can reject (N) or accept (Y) the offer. In period 1, after Player 1 rejects an offer, Player 2 can opt out (Q), or continue bargaining (C).

out only after *Player 1* rejects an offer. A similar analysis applies also to the model in which Player 2 can opt out both when responding to an offer and after Player 1 rejects an offer. We choose the case in which Player 2 is more restricted in order to simplify the analysis. The first two periods of the game we study are shown in Figure 3.6.

If $b < \delta^2/(1+\delta)$ then the outside option does not matter: the game has a unique subgame perfect equilibrium, which coincides with the subgame perfect equilibrium of the game in which Player 2 has no outside option. This corresponds to the first case in Proposition 3.5. We require $b < \delta^2/(1+\delta)$, rather than $b < \delta/(1+\delta)$ as in the model of the previous section in order that, if the players make offers and respond to offers as in the subgame perfect equilibrium of the game in which there is no outside option, then it is optimal for Player 2 to continue bargaining rather than opt out when Player 1 rejects an offer. (If Player 2 opts out then he collects b immediately. If he continues bargaining, then by accepting the agreement

$(1/(1+\delta), \delta/(1+\delta))$ that Player 1 proposes he can obtain $\delta/(1+\delta)$ with one period of delay, which is worth $\delta^2/(1+\delta)$ now.)

If $\delta^2/(1+\delta) \leq b \leq \delta^2$ then we obtain a result quite different from that in Proposition 3.5. There is a multiplicity of subgame perfect equilibria: for every $\xi \in [1-\delta, 1-b/\delta]$ there is a subgame perfect equilibrium that ends with immediate agreement on $(\xi, 1-\xi)$. In particular, there are equilibria in which Player 2 receives a payoff that *exceeds* the value of his outside option. In these equilibria Player 2 uses his outside option as a credible threat. Note that for this range of values of b we do not fully characterize the set of subgame perfect equilibria, although we do show that the presence of the outside option does not harm Player 2.

Proposition 3.6 *Consider the bargaining game described above, in which Player 2 can opt out only after Player 1 rejects an offer, as in Figure 3.6. Assume that the players have time preferences with the same constant discount factor $\delta < 1$, and that their payoffs in the event that Player 2 opts out in period t are $(0, \delta^t b)$, where $b < 1$.*

1. *If $b < \delta^2/(1+\delta)$ then the game has a unique subgame perfect equilibrium, which coincides with the subgame perfect equilibrium of the game in which Player 2 has no outside option. That is, Player 1 always proposes the agreement $(1/(1+\delta), \delta/(1+\delta))$ and accepts any proposal y in which $y_1 \geq \delta/(1+\delta)$, and Player 2 always proposes the agreement $(\delta/(1+\delta), 1/(1+\delta))$, accepts any proposal x in which $x_2 \geq \delta/(1+\delta)$, and never opts out. The outcome is that agreement is reached immediately on $(1/(1+\delta), \delta/(1+\delta))$.*
2. *If $\delta^2/(1+\delta) \leq b \leq \delta^2$ then there are many subgame perfect equilibria. In particular, for every $\xi \in [1-\delta, 1-b/\delta]$ there is a subgame perfect equilibrium that ends with immediate agreement on $(\xi, 1-\xi)$. In every subgame perfect equilibrium Player 2's payoff is at least $\delta/(1+\delta)$.*

Proof. We prove each part separately.

1. First consider the case $b < \delta^2/(1+\delta)$. The result follows from Theorem 3.4 once we show that, in any SPE, after *every* history it is optimal for Player 2 to continue bargaining, rather than to opt out. Let M_1 and m_2 be defined as in the proof of Proposition 3.5. By the arguments in Steps 1 and 2 of the proof of Theorem 3.4 we have $m_2 \geq 1 - \delta M_1$ and $M_1 \leq 1 - \delta m_2$, so that $m_2 \geq 1/(1+\delta)$. Now consider Player 2's decision to opt out. If he does so he obtains b immediately. If he continues bargaining and rejects Player 1's offer, play moves into a subgame in which he is first to make an offer. In this subgame he obtains at least m_2. He receives this payoff with two periods of delay, so it is worth at least $\delta^2 m_2 \geq \delta^2/(1+\delta)$

		η^*	b/δ	EXIT
1	proposes	$(1-\eta^*, \eta^*)$	$(1-b/\delta, b/\delta)$	$(1-\delta, \delta)$
	accepts	$x_1 \geq \delta(1-\eta^*)$	$x_1 \geq \delta(1-b/\delta)$	$x_1 \geq 0$
2	proposes	$(\delta(1-\eta^*)\ ,\ 1-\delta(1-\eta^*))$	$(\delta(1-b/\delta)\ ,\ 1-\delta(1-b/\delta))$	$(0,1)$
	accepts	$x_2 \geq \eta^*$	$x_2 \geq b/\delta$	$x_2 \geq \delta$
	opts out?	no	no	yes
	Transitions	Go to EXIT if Player 1 proposes x with $x_1 > 1-\eta^*$.	Go to EXIT if Player 1 proposes x with $x_1 > 1-b/\delta$.	Go to b/δ if Player 2 continues bargaining after Player 1 rejects an offer.

Table 3.5 The subgame perfect equilibrium in the proof of Part 2 of Proposition 3.6.

to him. Thus, since $b < \delta^2/(1+\delta)$, after any history it is better for Player 2 to continue bargaining than to opt out.

2. Now consider the case $\delta/(1+\delta) \leq b \leq \delta^2$. As in Part 1, we have $m_2 \geq 1/(1+\delta)$. We now show that for each $\eta^* \in [b/\delta, \delta]$ there is an SPE in which Player 2's utility is η^*. Having done so, we use these SPEs to show that for any $\xi^* \in [\delta b, \delta]$ there is an SPE in which Player 2's payoff is ξ^*. Since Player 2 can guarantee himself a payoff of δb by rejecting every offer of Player 1 in the first period and opting out in the second period, there is no SPE in which his payoff is less than δb. Further, since Player 2 must accept any offer x in which $x_2 > \delta$ in period 0 there is clearly no SPE in which his payoff exceeds δ. Thus our arguments show that the set of payoffs Player 2 obtains in SPEs is precisely $[\delta b, b]$.

Let $\eta^* \in [b/\delta, \delta]$. An SPE is given in Table 3.5. (For a discussion of this method of representing an equilibrium, see Section 3.5. Note that, as always, the initial state is the one in the leftmost column, and the transitions between states occur immediately after the events that trigger them.)

We now argue that this pair of strategies is an SPE. The analysis of the optimality of Player 1's strategy is straightforward. Consider Player 2. Suppose that the state is $\eta \in \{b/\delta, \eta^*\}$ and Player 1 proposes an agreement x with $x_1 \leq 1-\eta$. If Player 2 accepts this offer, as he is supposed to, he obtains the payoff $x_2 \geq \eta$. If he rejects the offer, then the state remains

η, and, given Player 1's strategy, the best action for Player 2 is either to propose the agreement y with $y_1 = \delta(1-\eta)$, which Player 1 accepts, or to propose an agreement that Player 1 rejects and opt out. The first outcome is worth $\delta[1-\delta(1-\eta)]$ to Player 2 today, which, under our assumption that $\eta^* \geq b/\delta \geq \delta/(1+\delta)$, is equal to at most η. The second outcome is worth $\delta b < b/\delta \leq \eta^*$ to Player 2 today. Thus it is optimal for Player 2 to accept the offer x. Now suppose that Player 1 proposes an agreement x in which $x_1 > 1-\eta\,(\geq 1-\delta)$. Then the state changes to EXIT. If Player 2 accepts the offer then he obtains $x_2 < \eta \leq \delta$. If he rejects the offer then by proposing the agreement $(0,1)$ he can obtain δ. Thus it is optimal for him to reject the offer x.

Now consider the choice of Player 2 after Player 1 has rejected an offer. Suppose that the state is η. If Player 2 opts out, then he obtains b. If he continues bargaining then by accepting Player 1's offer he can obtain η with one period of delay, which is worth $\delta\eta \geq b$ now. Thus it is optimal for Player 2 to continue bargaining.

Finally, consider the behavior of Player 2 in the state EXIT. The analysis of his acceptance and proposal policies is straightforward. Consider his decision when Player 1 rejects an offer. If he opts out then he obtains b immediately. If he continues bargaining then the state changes to b/δ, and the best that can happen is that he accepts Player 1's offer, giving him a utility of b/δ with one period of delay. Thus it is optimal for him to opt out. □

If $\delta^2 < b < 1$ then there is a unique subgame perfect equilibrium, in which Player 1 always proposes $(1-\delta,\delta)$ and accepts any offer, and Player 2 always proposes $(0,1)$, accepts any offer x in which $x_2 \geq \delta$, and always opts out.

We now come back to a comparison of the models in this section and the previous one. There are two interesting properties of the equilibria. First, when the value b to Player 2 of the outside option is relatively low—lower than it is in the unique subgame perfect equilibrium of the game in which he has no outside option—then his threat to opt out is not credible, and the presence of the outside option does not affect the outcome. Second, when the value of b is relatively high, the execution of the outside option is a credible threat, from which Player 2 can gain. The models differ in the way that the threat can be translated into a bargaining advantage. Player 2's position is stronger in the second model than in the first. In the second model he can make an offer that, given his threat, is effectively a "take-it-or-leave-it" offer. In the first model Player 1 has the right to make the last offer before Player 2 exercises his threat, and therefore she can ensure that Player 2 not get more than b. We conclude that the existence

of an outside option for a player affects the outcome of the game only if its use is credible, and the extent to which it helps the player depends on the possibility of making a "take-it-or-leave-it" offer, which in turn depends on the bargaining procedure.

3.13 A Game of Alternating Offers with Three Bargainers

Here we consider the case in which *three* players have access to a "pie" of size 1 if they can agree how to split it between them. Agreement requires the approval of all three players; no subset can reach agreement. There are many ways of extending the bargaining game of alternating offers to this case. An extension that appears to be natural was suggested and analyzed by Shaked; it yields the disappointing result that if the players are sufficiently patient then for *every* partition of the pie there is a subgame perfect equilibrium in which immediate agreement is reached on that partition.

Shaked's game is the following. In the first period, Player 1 proposes a partition (i.e. a vector $x = (x_1, x_2, x_3)$ with $x_1 + x_2 + x_3 = 1$), and Players 2 and 3 in turn accept or reject this proposal. If either of them rejects it, then play passes to the next period, in which it is Player 2's turn to propose a partition, to which Players 3 and 1 in turn respond. If at least one of them rejects the proposal, then again play passes to the next period, in which Player 3 makes a proposal, and Players 1 and 2 respond. Players rotate proposals in this way until a proposal is accepted by both responders. The players' preferences satisfy A1 through A6 of Section 3.3. Recall that $v_i(x_i, t)$ is the present value to Player i of the agreement x in period t (see (3.1)).

Proposition 3.7 *Suppose that the players' preferences satisfy assumptions A1 through A6 of Section 3.3, and $v_i(1, 1) \geq 1/2$ for $i = 1, 2, 3$. Then for any partition x^* of the pie there is a subgame perfect equilibrium of the three-player bargaining game defined above in which the outcome is immediate agreement on the partition x^*.*

Proof. Fix a partition x^*. Table 3.6, in which e^i is the ith unit vector, describes a subgame perfect equilibrium in which the players agree on x^* immediately. (Refer to Section 3.5 for a discussion of our method for presenting equilibria.) In each state $y = (y_1, y_2, y_3)$, each Player i proposes the partition y and accepts the partition x if and only if $x_i \geq v_i(y_i, 1)$. If, in any state y, a player proposes an agreement x for which he gets more than y_i, then there is a transition to the state e^j, where $j \neq i$ is the player with the lowest index for whom $x_j < 1/2$. As always, any transition between states occurs immediately after the event that triggers it; that is, immediately after an offer is made, *before* the response. Note that whenever

		x^*	e^1	e^2	e^3
1	proposes	x^*	e^1	e^2	e^3
	accepts	$x_1 \geq v_1(x_1^*, 1)$	$x_1 \geq v_1(1, 1)$	$x_1 \geq 0$	$x_1 \geq 0$
2	proposes	x^*	e^1	e^2	e^3
	accepts	$x_2 \geq v_2(x_2^*, 1)$	$x_2 \geq 0$	$x_2 \geq v_2(1, 1)$	$x_2 \geq 0$
3	proposes	x^*	e^1	e^2	e^3
	accepts	$x_3 \geq v_3(x_3^*, 1)$	$x_3 \geq 0$	$x_3 \geq 0$	$x_3 \geq v_3(1, 1)$
Transitions		If, in any state y, any Player i proposes x with $x_i > y_i$, then go to state e^j, where $j \neq i$ is the player with the lowest index for whom $x_j < 1/2$.			

Table 3.6 A subgame perfect equilibrium of Shaked's three-player bargaining game. The players' preferences are assumed to be such that $v_i(1, 1) \geq 1/2$ for $i = 1, 2, 3$. The agreement x^* is arbitrary, and e^i denotes the ith unit vector.

Player i proposes an agreement x for which $x_i > 0$ there is at least one player j for whom $x_j < 1/2$.

To see that these strategies form a subgame perfect equilibrium, first consider Player i's rule for accepting offers. If, in state y, Player i has to respond to an offer, then the most that he can obtain if he rejects the offer is y_i with one period of delay, which is worth $v_i(y_i, 1)$ to him. Thus acceptance of x if and only if $x_i \geq v_i(y_i, 1)$ is a best response to the other players' strategies. Now consider Player i's rule for making offers in state y. If he proposes x with $x_i > y_i$ then the state changes to e^j, j rejects i's proposal (since $x_j < 1/2 \leq v_i(e_j^j, 1) = v_i(1, 1)$), and i receives 0. If he proposes x with $x_i \leq y_i$ then either this offer is accepted or it is rejected and Player i obtains at most y_i in the next period. Thus it is optimal for Player i to propose y. □

The main force holding together the equilibrium in this proof is that one of the players is "rewarded" for rejecting a deviant offer—after his rejection, he obtains all of the pie. The result stands in sharp contrast to Theorem 3.4, which shows that a two-player bargaining game of alternating offers has a unique subgame perfect equilibrium. The key difference between the two situations seems to be the following. When there are three (or more) players one of the responders can always be compensated for rejecting a deviant offer, while when there are only two players this is not so. For example, in the two-player game there is no subgame perfect equilibrium

in which Player 1 proposes an agreement x in which she obtains less than $1 - v_2(1,1)$, since if she deviates and proposes an agreement y for which $x_1 < y_1 < 1 - v_2(1,1)$, then Player 2 must accept this proposal (because he can obtain at most $v_2(1,1)$ by rejecting it).

Several routes may be taken in order to isolate a unique outcome in Shaked's three-player game. For example, it is clear that the only subgame perfect equilibrium in which the players' strategies are stationary has a form similar to the unique subgame perfect equilibrium of the two-player game. (If the players have time preferences with a common constant discount factor δ, then this equilibrium leads to the division $(\xi, \delta\xi, \delta^2\xi)$ of the pie, where $\xi(1 + \delta + \delta^2) = 1$.) However, the restriction to stationary strategies is extremely strong (see the discussion at the end of Section 3.4). A more appealing route is to modify the structure of the game. For example, Perry and Shaked have proposed a game in which the players rotate in making demands. Once a player has made a demand, he may not subsequently increase this demand. The game ends when the demands sum to at most one. At the moment, no complete analysis of this game is available.

Notes

Most of the material in this chapter is based on Rubinstein (1982). For a related presentation of the material, see Rubinstein (1987). The proof of Theorem 3.4 is a modification of the original proof in Rubinstein (1982), following Shaked and Sutton (1984a). The discussion in Section 3.10.3 of the effect of diminishing the amount of time between a rejection and a counteroffer is based on Binmore (1987a, Section 5); the model in which the proposer is chosen randomly at the beginning of each period is taken from Binmore (1987a, Section 10). The model in Section 3.12.1, in which a player can opt out of the game, was suggested by Binmore, Shaked, and Sutton; see Shaked and Sutton (1984b), Binmore (1985), and Binmore, Shaked, and Sutton (1989). It is further discussed in Sutton (1986). Section 3.12.2 is based on Shaked (1994). The modeling choice between a finite and an infinite horizon which is discussed in Section 3.11 is not peculiar to the field of bargaining theory. In the context of repeated games, Aumann (1959) expresses a view similar to the one here. For a more detailed discussion of the issue, see Rubinstein (1991). Proposition 3.7 is due to Shaked (see also Herrero (1984)).

The first to investigate the alternating offer procedure was Ståhl (1972, 1977). He studies subgame perfect equilibria by using backwards induction in finite horizon models. When the horizons in his models are infinite he postulates nonstationary time preferences, which lead to the existence of a "critical period" at which one player prefers to yield rather than to con-

tinue, independently of what might happen next. This creates a "last interesting period" from which one can start the backwards induction. (For further discussion, see Ståhl (1988).) Other early work is that of Krelle (1975, 1976, pp. 607–632), who studies a T-period model in which a firm and a worker bargain over the division of the constant stream of profit (1 unit each period). Until an agreement is reached, both parties obtain 0 each period. Krelle notices that in the unique subgame perfect equilibrium of his game the wage converges to 1/2 as T goes to infinity.

As an alternative to using subgame perfect equilibrium as the solution in the bargaining game of alternating offers, one can consider the set of strategy pairs which remain when dominated strategies are sequentially eliminated. (A player's strategy is *dominated* if the player has another strategy that yields him at least as high a payoff, whatever strategy the other player uses, and yields a higher payoff against at least one of the other player's strategies.)

Among the variations on the bargaining game of alternating offers that have been studied are the following. Binmore (1987b) investigates the consequences of relaxing the assumptions on preferences (including the assumption of stationarity). Muthoo (1991) and van Damme, Selten, and Winter (1990) analyze the case in which the set of agreements is finite. Perry and Reny (1993) (see also Sákovics (1993)) study a model in which time runs continuously and players choose when to make offers. An offer must stand for a given length of time, during which it cannot be revised. Agreement is reached when the two outstanding offers are compatible. In every subgame perfect equilibrium an agreement is accepted immediately, and this agreement lies between x^* and y^* (see (3.3)). Muthoo (1992) considers the case in which the players can commit at the beginning of the game not to accept certain offers; they can revoke this commitment later only at a cost. Muthoo (1990) studies a model in which each player can withdraw from an offer if his opponent accepts it; he shows that all partitions can be supported by subgame perfect equilibria in this case.

Haller (1991), Haller and Holden (1990), and Fernandez and Glazer (1991) (see also Jones and McKenna (1988)) study a situation in which a firm and a union bargain over the stream of surpluses. In any period in which an offer is rejected, the union has to decide whether to strike (in which case it obtains a fixed payoff) or not (in which case it obtains a given wage). The model has a great multiplicity of subgame perfect equilibria, including some in which there is a delay, during which the union strikes, before an agreement is reached. This model is a special case of an interesting family of games in which in any period that an offer is rejected each bargainer has to choose an action from some set (see Okada (1991a, 1991b)). These

games interlace the structure of a repeated game with that of a bargaining game of alternating offers.

Admati and Perry (1991) study a model in which two players alternately contribute to a joint project which, upon completion, yields each of them a given payoff. Their model can be interpreted also as a variant of the bargaining game of alternating offers in which neither player can retreat from concessions he made in the past. Two further variants of the bargaining game of alternating offers, in the framework of a model of debt-renegotiation, are studied by Bulow and Rogoff (1989) and Fernandez and Rosenthal (1990).

The idea of endogenizing the timetable of bargaining when many issues are being negotiated is studied by Fershtman (1990) and Herrero (1988).

Models in which offers are made simultaneously are discussed, and compared with the model of alternating offers, by Chatterjee and Samuelson (1990), Stahl (1990), and Wagner (1984). Clemhout and Wan (1988) compare the model of alternating offers with a model of bargaining as a differential game (see also Leitmann (1973) and Fershtman (1989)).

Wolinsky (1987), Chikte and Deshmukh (1987), and Muthoo (1989) study models in which players may search for outside options while bargaining. For example, in Wolinsky's model both players choose the intensity with which to search for an outside option in any period in which there is disagreement; in Muthoo's model, one of the players may temporarily leave the bargaining table to search for an outside option.

Work on bargaining among more than two players includes the following. Haller (1986) points out that if the responses to an offer in a bargaining game of alternating offers with more than two players are simultaneous, rather than sequential, then the restriction on preferences in Proposition 3.7 is unnecessary. Jun (1987) and Chae and Yang (1988) study a model in which the players rotate in proposing a share for the next player in line; acceptance leads to the exit of the accepting player from the game. Various decision-making procedures in committees are studied by Dutta and Gevers (1984), Baron and Ferejohn (1987, 1989), and Harrington (1990). For example, Baron and Ferejohn (1989) compare a system in which in any period the committee members vote on a single proposal with a system in which, before a vote, any member may propose an amendment to the proposal under consideration. Chatterjee, Dutta, Ray, and Sengupta (1993) and Okada (1988b) analyze multi-player bargaining in the context of a general cooperative game, as do Harsanyi (1974, 1981) and Selten (1981), who draw upon semicooperative principles to narrow down the set of equilibria.

CHAPTER 4

The Relation between the Axiomatic and Strategic Approaches

4.1 Introduction

In Chapters 2 and 3 we took different approaches to the study of bargaining. The model in Chapter 2, due to Nash, is axiomatic: we start with a list of properties the solution is required to satisfy. By contrast, the model of alternating offers in Chapter 3 is strategic: we formulate the bargaining process as a specific extensive game. In this chapter we study the relation between the two approaches.

Nash's axiomatic model has advantages that are hard to exaggerate. It achieves great generality by avoiding any specification of the bargaining process; the solution defined by the axioms is unique, and its simple form is highly tractable, facilitating application. However, the axiomatic approach, and Nash's model in particular, has drawbacks. As we discussed in Chapter 2, it is difficult to assess how reasonable some axioms are without having in mind a specific bargaining procedure. In particular, Nash's axioms of Independence of Irrelevant Alternatives (IIA) and Pareto Efficiency (PAR) are hard to defend in the abstract. Further, within the axiomatic approach one cannot address issues relating directly to the bargaining process. For example, in Section 3.12 we used a strategic model to ask what is

the effect on the negotiated outcome of a player being able to terminate the negotiations. Nash's axiomatic model is powerless to analyze this question, which is perfectly suited for analysis within a strategic model.

Our investigation of the relation between the axiomatic and strategic approaches is intended to clarify the scope of the axiomatic approach. Unless we can find a sensible strategic model that has an equilibrium corresponding to the Nash solution, the appeal of Nash's axioms is in doubt. The characteristics of such a strategic model clarify the range of situations in which the axioms are reasonable.

The idea of relating axiomatic solutions to equilibria of strategic models was suggested by Nash (1953) and is now known as the "Nash program". In this chapter we pursue the Nash program by showing that there is a close connection between the Nash solution and the subgame perfect equilibrium outcome in the bargaining game of alternating offers we studied in Chapter 3. Also we show a connection between the Nash solution and the equilibria of a strategic model studied by Nash himself. These results reinforce Nash's claim that

> [t]he two approaches to the problem, via the negotiation model or via the axioms, are complementary; each helps to justify and clarify the other. (Nash (1953, p. 129))

In addition to providing a context within which an axiomatic model is appropriate, a formal connection between an axiomatic solution and the equilibrium of a strategic model is helpful in applications. When we use a model of bargaining within an economic context, we need to map the primitive elements of the bargaining model into the economic problem. Frequently there are several mappings that appear reasonable. For example, there may be several candidates for the disagreement point in Nash's model. A strategic model for which the Nash solution is an equilibrium can guide us to an appropriate modeling choice. We discuss the implications of our results along these lines in Section 4.6.

Before we can link the solutions of an axiomatic and a strategic model formally, we need to establish a common underlying model. The primitive elements in Nash's model are the set of outcomes (the set of agreements and the disagreement event) and the preferences of the players on lotteries over this set. In the model of alternating offers in Chapter 3 we are given the players' preferences over agreements reached at various points in time, rather than their preferences over uncertain outcomes. We begin (in Section 4.2) by introducing uncertainty into a bargaining game of alternating offers and assuming that the players are indifferent to the timing of an agreement. Specifically, after any offer is rejected there is a chance that the bargaining will terminate, and a "breakdown" event will occur. The probability that bargaining is interrupted in this way is fixed. (Note that

breakdown is exogenous; in contrast to the model in Section 3.12, neither player has any influence over the possibility of breakdown.) We show that the limit of the subgame perfect equilibria as this probability converges to zero corresponds to the Nash solution of an appropriately defined bargaining problem.

In Section 4.3 we discuss the strategic game suggested by Nash, in which uncertainty about the consequences of the players' actions also intervenes in the bargaining process. Once again, we show that the equilibria of the strategic game are closely related to the Nash solution of a bargaining problem.

In Section 4.4 we take a different tack: we redefine the Nash solution, using information about the players' time preferences rather than information about their attitudes toward risk. We consider a sequence of bargaining games of alternating offers in which the length of a period converges to zero. We show that the limit of the subgame perfect equilibrium outcomes of the games in such a sequence coincides with the modified Nash solution. In Section 4.5 we study a game in which the players are impatient *and* there is a positive probability that negotiations will break down after any offer is rejected. Finally, in Section 4.6, we discuss the implications of our analysis for applications.

4.2 A Model of Alternating Offers with a Risk of Breakdown

4.2.1 *The Game*

Here we study a strategic model of bargaining similar to the model of alternating offers in Chapter 3. As before, the set of possible agreements is

$$X = \{(x_1, x_2) \in \mathcal{R}^2 \colon x_1 + x_2 = 1 \text{ and } x_i \geq 0 \text{ for } i = 1, 2\}$$

(the set of divisions of the unit pie), and the players alternately propose members of X. The game differs in two respects from the one we studied in Chapter 3. First, at the end of each period, after an offer has been rejected, there is a chance that the negotiation ends with the breakdown event B. Precisely, this event occurs independently with (exogenous) probability $0 < q < 1$ at the end of each period. Second, each player is indifferent about the period in which an agreement is reached. We denote the resulting extensive game by $\Gamma(q)$; the first two periods of the game are shown in Figure 4.1. We study the connection between the Nash solution and the limit of the subgame perfect equilibria of $\Gamma(q)$ as the probability q of breakdown becomes vanishingly small.

The possibility of breakdown is exogenous in the game $\Gamma(q)$. The risk of breakdown, rather than the players' impatience (as in Chapter 3), is the

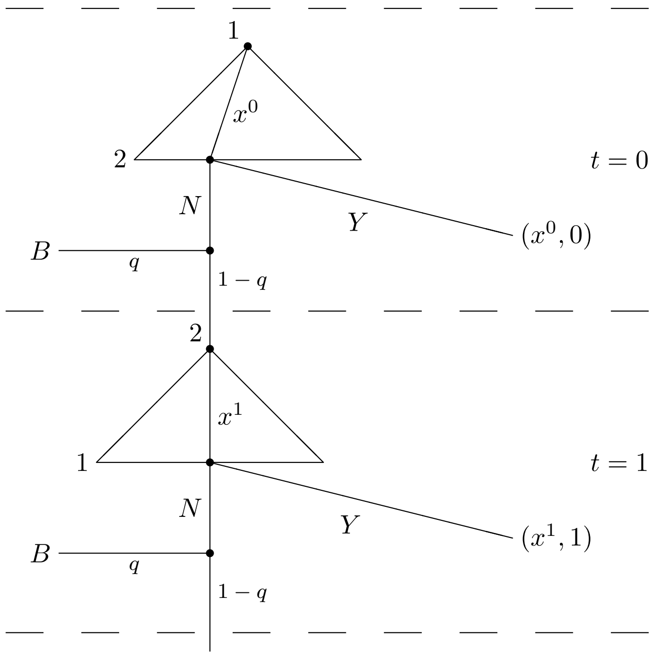


Figure 4.1 The first two periods of the bargaining game $\Gamma(q)$. After an offer is rejected, there is a probability q that negotiations break down, in which case the outcome B occurs.

basic force that motivates the players to reach an agreement as soon as possible. We can interpret a breakdown as the result of the intervention of a third party, who exploits the mutual gains. A breakdown can be interpreted also as the event that a threat made by one of the parties to halt the negotiations is actually realized. This possibility is especially relevant when a bargainer is a team (e.g. government), the leaders of which may find themselves unavoidably trapped by their own threats.

A strategy for each player in $\Gamma(q)$ is defined exactly as for a bargaining game of alternating offers (see Section 3.4). Let (σ, τ) be a pair of strategies that leads to the outcome (x, t) in a bargaining game of alternating offers (in which there is no possibility of breakdown). In the game $\Gamma(q)$ the probability that negotiation breaks down in any period is q, so that (σ, τ) leads to (x, t) with probability $(1-q)^t$ and to B with probability $1-(1-q)^t$.

Each player is indifferent to the timing of an outcome, so the period in which breakdown occurs is irrelevant to him. He is concerned only with the nature of the agreement that may be reached and the probability with

which this event occurs. Thus the consequence of a strategy pair that is relevant to a player's choice is the lottery in which some agreement x occurs with probability $(1-q)^t$, and the breakdown event B occurs with probability $1-(1-q)^t$. The probability q and the breakdown event B are fixed throughout, so this lottery depends only on the two variables x and t. We denote the lottery by $\langle\langle x, t\rangle\rangle$. Thus an outcome in $\Gamma(q)$, like an outcome in the bargaining game of alternating offers studied in Chapter 3, is a pair consisting of an agreement x, and a time t. The interpretations of the pairs (x, t) and $\langle\langle x, t\rangle\rangle$ are quite different. The first means that the agreement x is reached in period t, while the second is shorthand for the lottery in which x occurs with probability $(1-q)^t$, and B occurs with probability $1-(1-q)^t$. Our use of different delimiters for the outcomes (x, t) and $\langle\langle x, t\rangle\rangle$ serves as a reminder of the disparate interpretations.

However, a key element in the analysis of $\Gamma(q)$ is the exact correspondence between $\Gamma(q)$ and a bargaining game of alternating offers. Precisely, a pair of strategies that generates the outcome (x, t) in a bargaining game of alternating offers generates the outcome $\langle\langle x, t\rangle\rangle$ in the game $\Gamma(q)$; the pair of strategies that generates the outcome D (perpetual disagreement) in a bargaining game of alternating offers generates (with probability one) the outcome B in the game $\Gamma(q)$.

4.2.2 Preferences

In order to complete our description of the game $\Gamma(q)$, we need to specify the players' preferences over outcomes. We assume that each Player $i = 1, 2$ has a complete transitive reflexive preference ordering $\succsim_i$ over lotteries on $X \cup \{B\}$ that satisfies the assumptions of von Neumann and Morgenstern. Each preference ordering can thus be represented by the expected value of a continuous utility function $u_i\colon X \cup \{B\} \to \mathcal{R}$, which is unique up to an affine transformation. We assume that these utility functions satisfy the following three conditions, which are sufficient to guarantee that we can apply both the Nash solution and Theorem 3.4 to the game $\Gamma(q)$.

B1 (*Pie is desirable*) For any $x \in X$ and $y \in X$ we have $x \succ_i y$ if and only if $x_i > y_i$, for $i = 1, 2$.

B2 (*Breakdown is the worst outcome*) $(0, 1) \sim_1 B$ and $(1, 0) \sim_2 B$.

B3 (*Risk aversion*) For any $x \in X$, $y \in X$, and $\alpha \in [0, 1]$, each Player $i = 1, 2$ either prefers the certain outcome $\alpha x + (1-\alpha)y \in X$ to the lottery in which the outcome is x with probability α, and y with probability $1-\alpha$, or is indifferent between the two.

Under assumption B1, Player i's utility for $x \in X$ depends only on x_i, so we subsequently write $u_i(x_i)$ rather than $u_i(x_1, x_2)$. The significance of B2 is that there exists an agreement that both players prefer to B. The analysis can be easily modified to deal with the case in which some agreements are worse for one of the players than B: the set X has merely to be redefined to exclude such agreements. Without loss of generality, we set $u_i(B) = 0$ for $i = 1, 2$.

We now check that assumptions B1, B2, and B3 are sufficient to allow us to apply both the Nash solution and Theorem 3.4 to the game $\Gamma(q)$. First we check that the assumptions are sufficient to allow us to fit a bargaining problem to the game. Define

$$S = \{(s_1, s_2) \in \mathcal{R}^2 \colon (s_1, s_2) = (u_1(x_1), u_2(x_2)) \text{ for some } x \in X\}, \qquad (4.1)$$

and $d = (u_1(B), u_2(B)) = (0, 0)$. In order for $\langle S, d\rangle$ to be a bargaining problem (see Section 2.6.3), we need S to be the graph of a nonincreasing concave function and there to exist $s \in S$ for which $s_i > d_i$ for $i = 1, 2$. The first condition is satisfied because B1 and B3 imply that each u_i is increasing and concave. The second condition follows from B1 and B2.

Next we check that we can apply Theorem 3.4 to $\Gamma(q)$. To do so, we need to ensure that the preferences over lotteries of the form $\langle\!\langle x, t\rangle\!\rangle$ induced by the orderings $\succsim_i$ over lotteries on $X \cup \{B\}$ satisfy assumptions A1 through A6 of Section 3.3, when we replace the symbol (x, t) with $\langle\!\langle x, t\rangle\!\rangle$, and the symbol D by B. Under the assumptions above, each preference ordering over outcomes $\langle\!\langle x, t\rangle\!\rangle$ is complete and transitive, and

$$\langle\!\langle x, t\rangle\!\rangle \succ_i \langle\!\langle y, s\rangle\!\rangle \text{ if and only if } (1-q)^t u_i(x_i) > (1-q)^s u_i(y_i)$$

(since $u_i(B) = 0$).

It follows from B1 and B2 that $\langle\!\langle x, t\rangle\!\rangle \succsim_i B$ for all outcomes $\langle\!\langle x, t\rangle\!\rangle$, so that A1 is satisfied. From B1 we deduce that $\langle\!\langle x, t\rangle\!\rangle \succ_i \langle\!\langle y, t\rangle\!\rangle$ if and only if $x_i > y_i$, so that A2 is satisfied. Also $\langle\!\langle x, t\rangle\!\rangle \succsim_i \langle\!\langle x, s\rangle\!\rangle$ if $t < s$, with strict preference if $x_i > 0$ (since $u_i(x_i)$ is then positive by B1 and B2), so that A3 is satisfied. The continuity of each u_i ensures that A4 is satisfied, and A5 follows immediately. Finally, we show that A6 is satisfied. The continuity of u_i implies that for every $x \in X$ there exists $y \in X$ such that $u_i(y_i) = (1-q)^t u_i(x_i)$, so that $\langle\!\langle y, 0\rangle\!\rangle \sim_i \langle\!\langle x, t\rangle\!\rangle$. Hence the present value $v_i(x_i, 1)$ of the lottery $\langle\!\langle x, 1\rangle\!\rangle$ satisfies $u_i\left(v_i(x_i, 1)\right) = (1-q)u_i(x_i)$, or $u_i(x_i) - u_i\left(v_i(x_i, 1)\right) = qu_i(x_i)$. Let $x_i < y_i$. The concavity of u_i implies that

$$\frac{u_i(x_i) - u_i\left(v_i(x_i, 1)\right)}{x_i - v_i(x_i, 1)} \geq \frac{u_i(y_i) - u_i\left(v_i(y_i, 1)\right)}{y_i - v_i(y_i, 1)}.$$

Thus

$$\frac{qu_i(x_i)}{x_i - v_i(x_i, 1)} \geq \frac{qu_i(y_i)}{y_i - v_i(y_i, 1)}.$$

Since $u_i(x_i) < u_i(y_i)$ it follows that $x_i - v_i(x_i, 1) < y_i - v_i(y_i, 1)$, so that A6 is satisfied.

4.2.3 Subgame Perfect Equilibrium

Given that the players' preferences over lotteries of the form $\langle\langle x, t\rangle\rangle$ satisfy assumptions A1 through A6 of Section 3.3, we can deduce from Theorem 3.4 the character of the unique subgame perfect equilibrium of $\Gamma(q)$, for any fixed $q \in (0, 1)$. As we noted above, for every lottery $\langle\langle x, t\rangle\rangle$ there is an agreement $y \in X$ such that $\langle\langle y, 0\rangle\rangle \sim_i \langle\langle x, t\rangle\rangle$. Let $(x^*(q), y^*(q))$ be the unique pair of agreements satisfying

$$\langle\langle y^*(q), 0\rangle\rangle \sim_1 \langle\langle x^*(q), 1\rangle\rangle \qquad \text{and} \qquad \langle\langle x^*(q), 0\rangle\rangle \sim_2 \langle\langle y^*(q), 1\rangle\rangle$$

(see (3.4)). Transforming this into a statement about utilities, we have

$$u_1\left(y_1^*(q)\right) = (1-q)u_1\left(x_1^*(q)\right) \quad \text{and} \quad u_2\left(x_2^*(q)\right) = (1-q)u_2\left(y_2^*(q)\right). \tag{4.2}$$

Thus by Theorem 3.4 we have the following.

Proposition 4.1 *For each $q \in (0, 1)$ the game $\Gamma(q)$ has a unique subgame perfect equilibrium. In this equilibrium Player 1 proposes the agreement $x^*(q)$ in period 0, which Player 2 accepts.*

4.2.4 The Relation with the Nash Solution

We now show that there is a very close relation between the Nash solution of the bargaining problem $\langle S, d\rangle$, where S is defined in (4.1) and $d = (0, 0)$, and the limit of the unique subgame perfect equilibrium of $\Gamma(q)$ as $q \to 0$.

Proposition 4.2 *The limit, as $q \to 0$, of the agreement $x^*(q)$ reached in the unique subgame perfect equilibrium of $\Gamma(q)$ is the agreement given by the Nash solution of the bargaining problem $\langle S, d\rangle$, where S is defined in (4.1) and $d = (0, 0)$.*

Proof. It follows from (4.2) that $u_1\left(x_1^*(q)\right) u_2\left(x_2^*(q)\right) = u_1\left(y_1^*(q)\right) u_2\left(y_2^*(q)\right)$, and that $\lim_{q\to 0}\left[u_i\left(x_i^*(q)\right) - u_i\left(y_i^*(q)\right)\right] = 0$ for $i = 1, 2$. Thus $x^*(q)$ converges to the maximizer of $u_1(x_1)u_2(x_2)$ over S (see Figure 4.2). □

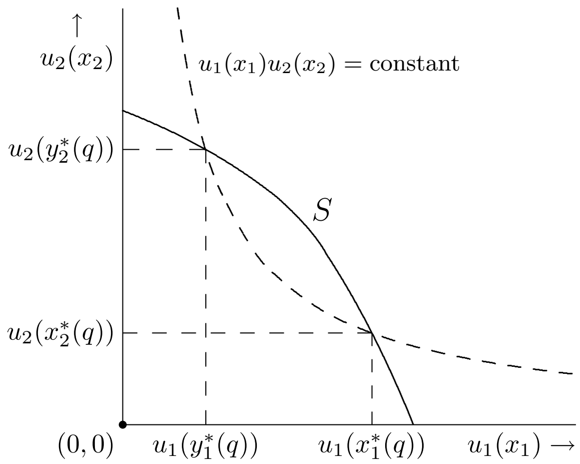


Figure 4.2 An illustration of the proof of Proposition 4.2.

This result is illustrated in Figure 4.3. It shows that if we perturb a bargaining game of alternating offers by introducing a small exogenous probability of breakdown then, when the players are indifferent to the timing of an agreement, the unique subgame perfect equilibrium outcome is close to the Nash solution of the appropriately defined bargaining problem. We discuss the implications of this result for applications of the Nash bargaining solution in Section 4.6.

4.3 A Model of Simultaneous Offers: Nash's "Demand Game"

Nash himself (1953) considered a strategic model of bargaining that "supports" his axiomatic solution. In this model, time plays no role. Although the model is static rather than sequential, and thus is a diversion from the main theme of the book, we present it here because of its central role in the development of the theory.

The game consists of a single stage, in which the two players simultaneously announce "demands". If these are compatible, then each player receives the amount he demanded; otherwise the disagreement event occurs. This game, like a bargaining game of alternating offers, has a plethoraof Nash equilibria. Moreover, the notion of subgame perfect equilibrium obviously has no power to discriminate among the equilibria, as it does in a bargaining game of alternating offers, since the game has no proper sub-

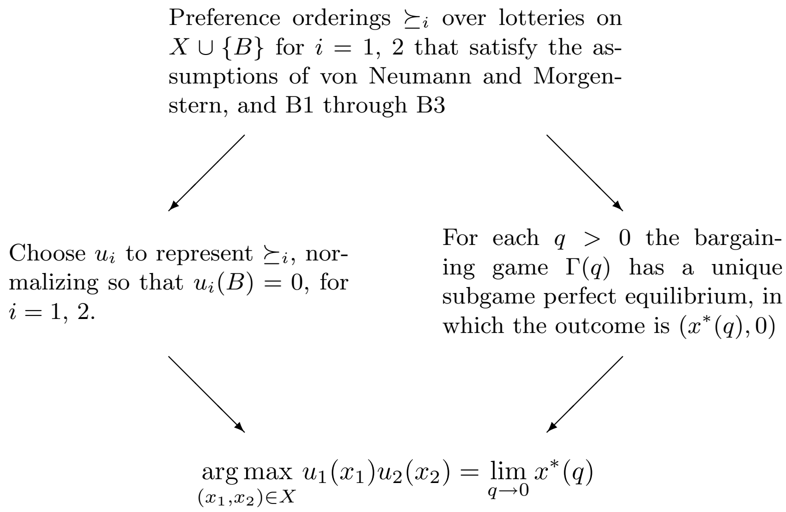


Figure 4.3 An illustration of Proposition 4.2.

games. In order to facilitate a comparison of the strategic and axiomatic models, Nash used a different approach to refine the set of equilibria—an approach that foreshadows the notions of "perfection" deriving from Selten's (1975) work.

4.3.1 The Demand Game

Let $\langle S, d\rangle$ be a bargaining problem (see Definition 2.1) in which S has a nonempty interior. Without loss of generality, let $d = (0, 0)$. Nash's *Demand Game* is the two-player strategic game G defined as follows. The strategy set of each Player $i = 1, 2$ is $\mathcal{R}_+$; the payoff function $h_i\colon \mathcal{R}_+ \times \mathcal{R}_+ \to \mathcal{R}$ of i is defined by

$$h_i(\sigma_1, \sigma_2) = \begin{cases} 0 & \text{if } (\sigma_1, \sigma_2) \notin S \\ \sigma_i & \text{if } (\sigma_1, \sigma_2) \in S. \end{cases}$$

An interpretation is that each Player i in G may "demand" any utility σ_i at least equal to what he gets in the event of disagreement. If the demands are infeasible, then each player receives his disagreement utility; if they are feasible, then each player receives the amount he demands.

The set of Nash equilibria of G consists of the set of strategy pairs that are strongly Pareto efficient and some strategy pairs (for example, those in

which each player demands more than the maximum he can obtain at any point in S) that yield the disagreement utility pair $(0, 0)$.

4.3.2 The Perturbed Demand Game

Given that the notion of Nash equilibrium puts so few restrictions on the nature of the outcome of a Demand Game, Nash considered a more discriminating notion of equilibrium, which is related to Selten's (1975) "perfect equilibrium". The idea is to require that an equilibrium be robust to perturbations in the structure of the game. There are many ways of formulating such a condition. We might, for example, consider a Nash equilibrium σ^* of a game Γ to be robust if every game in which the payoff functions are close to those of Γ has an equilibrium close to σ^*. Nash's approach is along these lines, though instead of requiring robustness to *all* perturbations of the payoff functions, Nash considered a *specific* class of perturbations of the payoff function, tailored to the interpretation of the Demand Game.

Precisely, perturb the Demand Game, so that there is some uncertainty in the neighborhood of the boundary of S. Suppose that if a pair of demands $(\sigma_1, \sigma_2) \in S$ is close to the boundary of S then, despite the compatibility of these demands, there is a positive probability that the outcome is the disagreement point d, rather than the agreement (σ_1, σ_2). Specifically, suppose that any pair of demands $(\sigma_1, \sigma_2) \in \mathcal{R}^2_+$ results in the agreement (σ_1, σ_2) with probability $P(\sigma_1, \sigma_2)$, and in the disagreement event with probability $1 - P(\sigma_1, \sigma_2)$. If $(\sigma_1, \sigma_2) \notin S$ then $P(\sigma_1, \sigma_2) = 0$ (incompatible demands cannot be realized); otherwise, $0 \le P(\sigma_1, \sigma_2) \le 1$, and $P(\sigma_1, \sigma_2) > 0$ for all (σ_1, σ_2) in the interior of[1] S. The payoff function of Player i $(= 1, 2)$ in the perturbed game is

$$h_i(\sigma_1, \sigma_2) = \sigma_i P(\sigma_1, \sigma_2). \tag{4.3}$$

We assume that the function $P: \mathcal{R}^2_+ \to [0, 1]$ defining the probability of breakdown in the perturbed game is differentiable. We further assume that P is quasi-concave, so that for each $\rho \in [0, 1]$ the set

$$\overline{P}(\rho) = \{(\sigma_1, \sigma_2) \in \mathcal{R}^2_+ : P(\sigma_1, \sigma_2) \ge \rho\} \tag{4.4}$$

is convex. (Note that this is consistent with the convexity of S.) A bargaining problem $\langle S, d\rangle$ in which $d = (0, 0)$, and a perturbing function P define a *Perturbed Demand Game* in which the strategy set of each player is $\mathcal{R}_+$ and the payoff function h_i of $i = 1, 2$ is defined in (4.3).

[1] Nash (1953) considers a slightly different perturbation, in which the probability of agreement is one everywhere in S, and tapers off toward zero outside S. See van Damme (1987, Section 7.5) for a discussion of this case.

4.3.3 Nash Equilibria of the Perturbed Games: A Convergence Result

Every Perturbed Demand Game has equilibria that yield the disagreement event. (Consider, for example, any strategy pair in which each player demands more than the maximum he can obtain in any agreement.) However, as the next result shows, the set of equilibria that generate *agreement* with positive probability is relatively small and converges to the Nash solution of $\langle S, d\rangle$ as the Hausdorff distance between S and $\overline{P^n}(1)$ converges to zero—i.e. as the perturbed game approaches the original demand game. (The Hausdorff distance between the set S and $T \subset S$ is the maximum distance between a point in S and the closest point in T.)

Proposition 4.3 *Let G^n be the Perturbed Demand Game defined by $\langle S, d\rangle$ and P^n. Assume that the Hausdorff distance between S and the set $\overline{P^n}(1)$ associated with P^n converges to zero as $n \to \infty$. Then every game G^n has a Nash equilibrium in which agreement is reached with positive probability, and the limit as $n \to \infty$ of every sequence $\{\sigma^{*n}\}_{n=1}^{\infty}$ in which σ^{*n} is such a Nash equilibrium is the Nash solution of $\langle S, d\rangle$.*

Proof. First we show that every perturbed game G^n has a Nash equilibrium in which agreement is reached with positive probability. Consider the problem

$$\max_{(\sigma_1,\sigma_2)\in\mathcal{R}^2_+} \sigma_1\sigma_2 P^n(\sigma_1,\sigma_2).$$

Since P^n is continuous, and equal to zero outside the compact set S, this problem has a solution $(\hat{\sigma}_1, \hat{\sigma}_2) \in S$. Further, since $P^n(\sigma_1,\sigma_2) > 0$ whenever (σ_1,σ_2) is in the interior of S, we have $\hat{\sigma}_i > 0$ for $i = 1$, 2 and $P^n(\hat{\sigma}_1,\hat{\sigma}_2) > 0$. Consequently $\hat{\sigma}_1$ maximizes $\sigma_1 P^n(\sigma_1,\hat{\sigma}_2)$ over $\sigma_1 \in \mathcal{R}_+$, and $\hat{\sigma}_2$ maximizes $\sigma_2 P^n(\hat{\sigma}_1,\sigma_2)$ over $\sigma_2 \in \mathcal{R}_+$. Hence $(\hat{\sigma}_1,\hat{\sigma}_2)$ is a Nash equilibrium of G^n.

Now let $(\sigma_1^*, \sigma_2^*) \in S$ be an equilibrium of G^n in which agreement is reached with positive probability. If $\sigma_i^* = 0$ then by the continuity of P^n, Player i can increase his demand and obtain a positive payoff. Hence $\sigma_i^* > 0$ for $i = 1$, 2. Thus by the assumption that P^n is differentiable, the fact that σ_i^* maximizes i's payoff given σ_j^* implies that[2]

$$\sigma_i^* D_i P^n(\sigma_1^*,\sigma_2^*) + P^n(\sigma_1^*,\sigma_2^*) = 0 \text{ for } i = 1, 2,$$

and hence

$$\frac{D_1 P^n(\sigma_1^*,\sigma_2^*)}{D_2 P^n(\sigma_1^*,\sigma_2^*)} = \frac{\sigma_2^*}{\sigma_1^*}. \tag{4.5}$$

Let $\pi^* = P^n(\sigma_1^*,\sigma_2^*)$, so that $(\sigma_1^*,\sigma_2^*) \in \overline{P^n}(\pi^*)$. The fact that (σ_1^*,σ_2^*) is a Nash equilibrium implies in addition that (σ_1^*,σ_2^*) is on the Pareto frontier

[2] We use $D_i f$ to denote the partial derivative of f with respect to its ith argument.

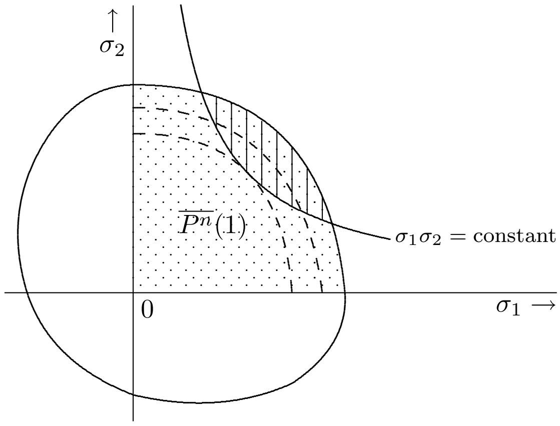


Figure 4.4 The Perturbed Demand Game. The area enclosed by the solid line is S. The dashed lines are contours of P^n. Every Nash equilibrium of the perturbed game in which agreement is reached with positive probability lies in the area shaded by vertical lines.

of $\overline{P^n}(\pi^*)$. It follows from (4.5) and the fact that P^n is quasi-concave that (σ_1^*, σ_2^*) is the maximizer of $\sigma_1\sigma_2$ subject to $P^n(\sigma_1, \sigma_2) \geq \pi^*$. In particular,

$$\sigma_1^*\sigma_2^* \geq \max_{(\sigma_1,\sigma_2)} \{\sigma_1\sigma_2 \colon (\sigma_1, \sigma_2) \in \overline{P^n}(1)\},$$

so that (σ_1^*, σ_2^*) lies in the shaded area of Figure 4.4. As $n \to \infty$, the set $\overline{P^n}(1)$ converges (in Hausdorff distance) to $S \cap \mathcal{R}^2_+$, so that this area converges to the Nash solution of $\langle S, d\rangle$.

Thus the limit of every sequence $\{\sigma^{*n}\}_{n=1}^{\infty}$ for which σ^{*n} is a Nash equilibrium of G^n and $P^n(\sigma^{*n}) > 0$ is the Nash solution of $\langle S, d\rangle$. □

The assumption that the perturbing functions P^n are differentiable is essential to the result. If not, then the perturbed games G^n may have Nash equilibria far from the Nash solution of $\langle S, d\rangle$, even when $\overline{P^n}(1)$ is very close to S.[3]

[3] Suppose, for example, that the intersection of the set S of agreement utilities with the nonnegative quadrant is the convex hull of $(0,0)$, $(1,0)$, and $(0,1)$ (the "unit simplex"), and define P^n on the unit simplex by

$$P^n(\sigma_1, \sigma_2) = \begin{cases} 1 & \text{if } 0 \leq \sigma_1 + \sigma_2 \leq 1 - 1/n \\ n(1 - \sigma_1 - \sigma_2) & \text{if } 1 - 1/n \leq \sigma_1 + \sigma_2 \leq 1. \end{cases}$$

Then any pair (σ_1, σ_2) in the unit simplex with $\sigma_1 + \sigma_2 = 1 - 1/n$ and $\sigma_i \geq 1/n$ for $i = 1, 2$ is a Nash equilibrium of G^n. Thus *all* points in the unit simplex that are on the Pareto frontier of S are limits of Nash equilibria of G^n.

The result provides additional support for the Nash solution. In a model, like that of the previous section, where some small amount of exogenous uncertainty interferes with the bargaining process, we have shown that all equilibria that lead to agreement with positive probability are close to the Nash solution of the associated bargaining problem. The result is different than that of the previous section in three respects. First, the demand game is static. Second, the disagreement point is always an equilibrium outcome of a perturbed demand game—the result restricts the character only of equilibria that result in agreement with positive probability. Third, the result depends on the differentiability and quasi-concavity of the perturbing function, characteristics that do not appear to be natural.

4.4 Time Preference

We now turn back to the bargaining model of alternating offers studied in Chapter 3, in which the players' impatience is the driving force. In this section we think of a period in the bargaining game as an interval of real time of length $\Delta > 0$, and examine the limit of the subgame perfect equilibria of the game as Δ approaches zero. Thus we generalize the discussion in Section 3.10.3, which deals only with time preferences with a constant discount rate.

We show that the limit of the subgame perfect equilibria of the bargaining game as the delay between offers approaches zero can be calculated using a simple formula closely related to the one used to characterize the Nash solution. However, we do not consider the limit to be the Nash solution, since the utility functions that appear in the formula reflect the players' time preferences, not their attitudes toward risk as in the Nash bargaining solution.

4.4.1 Bargaining Games with Short Periods

Consider a bargaining game of alternating offers (see Definition 3.1) in which the delay between offers is Δ: offers can be made only at a time in the denumerable set $\{0, \Delta, 2\Delta, \ldots\}$. We denote such a game by $\Gamma(\Delta)$. We wish to study the effect of letting Δ converge to zero. Since we want to allow any value of Δ, we start with a preference ordering for each player defined on the set $(X \times T_\infty) \cup \{D\}$, where $T_\infty = [0, \infty)$. For each $\Delta > 0$, such an ordering induces an ordering over the set $(X \times \{0, \Delta, 2\Delta, \ldots\}) \cup \{D\}$. In order to apply the results of Chapter 3, we impose conditions on the orderings over $(X \times T_\infty) \cup \{D\}$ so that the induced orderings satisfy conditions A1 through A6 of that chapter.

We require that each Player $i = 1, 2$ have a complete transitive reflexive preference ordering $\succsim_i$ over $(X \times T_\infty) \cup \{D\}$ that satisfies analogs of assumptions A1 through A6 in Chapter 3. Specifically, we assume that $\succsim_i$ satisfies the following.

C1 (*Disagreement is the worst outcome*) For every $(x,t) \in X \times T_\infty$ we have $(x,t) \succsim_i D$.

C2 (*Pie is desirable*) For any $t \in T_\infty$, $x \in X$, and $y \in X$ we have $(x,t) \succ_i (y,t)$ if and only if $x_i > y_i$.

We slightly strengthen A3 of Chapter 3 to require that each Player i be indifferent about the timing of an agreement x in which $x_i = 0$. This condition is satisfied by preferences with constant discount rates, but not for preferences with a constant cost of delay (see Section 3.3.3).

C3 (*Time is valuable*) For any $t \in T_\infty$, $s \in T_\infty$, and $x \in X$ with $t < s$ we have $(x,t) \succ_i (x,s)$ if $x_i > 0$, and $(x,t) \sim_i (x,s)$ if $x_i = 0$.

Assumptions A4 and A5 remain essentially unchanged.

C4 (*Continuity*) Let $\{(x_n,t_n)\}_{n=1}^\infty$ and $\{(y_n,s_n)\}_{n=1}^\infty$ be convergent sequences of members of $X \times T_\infty$ with limits (x,t) and (y,s), respectively. Then $(x,t) \succsim_i (y,s)$ whenever $(x_n,t_n) \succsim_i (y_n,s_n)$ for all n.

C5 (*Stationarity*) For any $t \in T_\infty$, $x \in X$, $y \in X$, and $\theta \geq 0$ we have $(x,t) \succ_i (y,t+\theta)$ if and only if $(x,0) \succ_i (y,\theta)$.

The fact that C3 is stronger than A3 allows us to deduce that for any outcome $(x,t) \in X \times T_\infty$ there exists an agreement $y \in X$ such that $(y,0) \sim_i (x,t)$. The reason is that by C3 and C2 we have $(x,0) \succsim_i (x,t) \succsim_i (z,t) \sim_i (z,0)$, where z is the agreement for which $z_i = 0$; the claim follows from C4. Consequently the present value $v_i(x_i,t)$ of an outcome (x,t) satisfies

$$(y,0) \sim_i (x,t) \text{ whenever } y_i = v_i(x_i,t) \tag{4.6}$$

(see (3.1) and (3.2)).

Finally, we strengthen A6. We require, in addition to A6, that the loss to delay be a *concave* function of the amount involved.

C6 (*Increasing and concave loss to delay*) The loss to delay $x_i - v_i(x_i,1)$ is an increasing and concave function of x_i.

The condition of convexity of v_i in x_i has no analog in the analysis of Chapter 3: it is an additional assumption we need to impose on preferences in order to obtain the result of this section. The condition is satisfied, for example, by time preferences with a constant discount rate, since the loss to delay in this case is linear.

4.4.2 Subgame Perfect Equilibrium

If the preference ordering $\succsim_i$ of Player i over $(X \times T_\infty) \cup \{D\}$ satisfies C1 through C6, then for any value of Δ the ordering induced over $(X \times \{0, \Delta, 2\Delta, \ldots\}) \cup \{D\}$ satisfies A1 through A6 of Chapter 3. Hence we can apply Theorem 3.4 to the game $\Gamma(\Delta)$. For any value of $\Delta > 0$, let $(x^*(\Delta), y^*(\Delta)) \in X \times X$ be the unique pair of agreements satisfying

$$(y^*(\Delta), 0) \sim_1 (x^*(\Delta), \Delta) \qquad \text{and} \qquad (x^*(\Delta), 0) \sim_2 (y^*(\Delta), \Delta)$$

(see (3.3) and (4.6)). We have the following.

Proposition 4.4 *Suppose that each player's preference ordering satisfies C1 through C6. Then for each $\Delta > 0$ the game $\Gamma(\Delta)$ has a unique subgame perfect equilibrium. In this equilibrium Player 1 proposes the agreement $x^*(\Delta)$ in period 0, which Player 2 accepts.*

4.4.3 The Relation with the Nash Solution

As we noted in the discussion after A5 on p. 34, preferences that satisfy A2 through A5 of Chapter 3 can be represented on $X \times T$ by a utility function of the form $\delta_i^t u_i(x_i)$. Under our stronger assumptions here we can be more specific. If the preference ordering $\succsim_i$ on $(X \times T_\infty) \cup \{D\}$ satisfies C1 through C6, then there exists $\delta_i \in (0, 1)$ such that for each $\delta_i \geq \overline{\delta}_i$ there is a increasing *concave* function $u_i\colon X \to \mathcal{R}$, unique up to multiplication by a positive constant, with the property that $\delta_i^t u_i(x_i)$ represents $\succsim_i$ on $X \times T_\infty$. (In the case that the set of times is discrete, this follows from Proposition 1 of Fishburn and Rubinstein (1982); the methods in the proof of their Theorem 2 can be used to show that the result holds also when the set of times is T_∞.)

Now suppose that $\delta_i^t u_i(x_i)$ represents $\succsim_i$ on $X \times T_\infty$, and $0 < \epsilon_i < 1$. Then $[\delta_i^t u_i(x_i)]^{(\log \epsilon_i)/(\log \delta_i)} = \epsilon_i^t [u_i(x_i)]^{(\log \epsilon_i)/(\log \delta_i)}$ also represents $\succsim_i$. We conclude that if in addition $\epsilon_i^t w_i(x_i)$ represents $\succsim_i$ then $w_i(x_i) = K_i [u_i(x_i)]^{(\log \epsilon_i)/(\log \delta_i)}$ for some $K_i > 0$.

We now consider the limit of the subgame perfect equilibrium outcome of $\Gamma(\Delta)$ as $\Delta \to 0$. Fix a *common* discount factor $\delta < 1$ that is large enough for there to exist increasing concave functions u_i ($i = 1$, 2) with

the property that $\delta^t u_i(x_i)$ represents $\succsim_i$. Let

$$S = \{s \in \mathcal{R}^2 : s = (u_1(x_1), u_2(x_2)) \text{ for some } (x_1, x_2) \in X\}, \tag{4.7}$$

and let $d = (0, 0)$. Since each u_i is increasing and concave, S is the graph of a nonincreasing concave function. Further, by the second part of C3 we have $u_i(0) = 0$ for $i = 1$, 2, so that by C2 there exists $s \in S$ such that $s_i > d_i$ for $i = 1$, 2. Thus $\langle S, d\rangle$ is a bargaining problem. The set S depends on the discount factor δ we chose. However, the Nash solution of $\langle S, d\rangle$ is independent of this choice: the maximizer of $u_1(x_1)u_2(x_2)$ is also the maximizer of $K_1K_2[u_1(x_1)u_2(x_2)]^{(\log\epsilon)/(\log\delta)}$ for any $0 < \epsilon < 1$.

We emphasize that in constructing the utility functions u_i for $i = 1$, 2, we use the *same* discount factor δ. In some contexts, the economics of a problem suggests that the players' preferences be represented by particular utility functions. These functions do not necessarily coincide with the functions that must be used to construct S. For example, suppose that in some problem it is natural for the players to have the utility functions $\delta_i^t x_i$ for $i = 1$, 2, where $\delta_1 > \delta_2$. Then the appropriate functions u_i are constructed as follows. Let $\delta = \delta_1$, and define u_1 by $u_1(x_1) = x_1$ and u_2 by $u_2(x_2) = x_2^{(\log\delta_1)/(\log\delta_2)}$ (*not* by $u_2(x_2) = x_2$).

The main result of this section is the following. It is illustrated in Figure 4.5.

Proposition 4.5 *If the preference ordering of each player satisfies C1 through C6, then the limit, as $\Delta \to 0$, of the agreement $x^*(\Delta)$ reached in the unique subgame perfect equilibrium of $\Gamma(\Delta)$ is the agreement given by the Nash solution of the bargaining problem $\langle S, d\rangle$, where S is defined in (4.7) and $d = (0, 0)$.*

Proof. It follows from Proposition 4.4 that $u_1(y_1^*(\Delta)) = \delta^\Delta u_1(x_1^*(\Delta))$ and $u_2(x_2^*(\Delta)) = \delta^\Delta u_2(y_2^*(\Delta))$. The remainder of the argument parallels that in the proof of Proposition 4.2. □

4.4.4 Symmetry and Asymmetry

Suppose that Player i's preferences in a bargaining game of alternating offers are represented by $\delta_i^t w_i(x_i)$, where w_i is concave ($i = 1$, 2), and $\delta_1 > \delta_2$. To find the limit, as the delay between offers converges to zero, of the subgame perfect equilibrium outcome of this game, we can use Proposition 4.5 as follows. Choose δ_1 to be the common discount factor with respect to which preferences are represented, and set $u_1 = w_1$. Let $u_2(x_2) = [w_2(x_2)]^{(\log\delta_1)/(\log\delta_2)}$, so that u_2 is increasing and concave, and

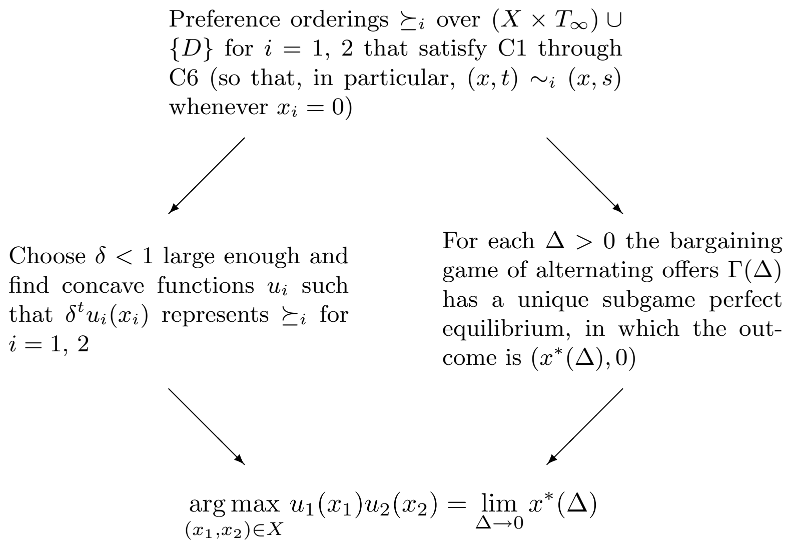


Figure 4.5 An illustration of Proposition 4.5.

$\delta_1^t u_2(x_2)$ represents Player 2's preferences. By Proposition 4.5 the limit of the agreement reached in a subgame perfect equilibrium of a bargaining game of alternating offers as the length of a period converges to zero is the Nash solution of $\langle S, d\rangle$, where S is defined in (4.7). This Nash solution is given by

$$\arg\max_{(x_1,x_2)\in X} u_1(x_1)u_2(x_2) = \arg\max_{(x_1,x_2)\in X} w_1(x_1)[w_2(x_2)]^{(\log\delta_1)/(\log\delta_2)}, \tag{4.8}$$

or alternatively

$$\arg\max_{(x_1,x_2)\in X} [w_1(x_1)]^{\alpha}[w_2(x_2)]^{1-\alpha},$$

where $\alpha = (\log\delta_2)/(\log\delta_1 + \log\delta_2)$. Thus the solution is an *asymmetric* Nash solution (see (2.4)) of the bargaining problem constructed using the *original* utility functions w_1 and w_2. The degree of asymmetry is determined by the disparity in the discount factors.

If the original utility function w_i of each Player i is linear ($w_i(x_i) = x_i$), we can be more specific. In this case, the agreement given by (4.8) is

$$\left(\frac{\log\delta_2}{\log\delta_1 + \log\delta_2}, \frac{\log\delta_1}{\log\delta_1 + \log\delta_2}\right),$$

which coincides (as it should!) with the result in Section 3.10.3.

In the case we have examined so far, the players are asymmetric because they value time differently. Another source of asymmetry may be embedded in the structure of the game: the amount of time that elapses between a rejection and an offer may be different for Player 1 than for Player 2. Specifically, consider a bargaining game of alternating offers $\Gamma(\gamma_1, \gamma_2)$, in which the time that elapses between a rejection and a counteroffer by Player i is $\gamma_i\Delta$ (= 1, 2). As Δ converges to zero, the length of time between any rejection and counteroffer diminishes, while the ratio of these times for Players 1 and 2 remains constant. Suppose that there is a common discount factor δ and a function u_i for each Player i such that his preferences are represented by $\delta^t u_i(x_i)$. The preferences induced over the outcomes (x, n), where n indexes the rounds of negotiation in $\Gamma(\gamma_1, \gamma_2)$, are not stationary. Nevertheless, as we noted in Section 3.10.4, the game $\Gamma(\gamma_1, \gamma_2)$ has a unique subgame perfect equilibrium; this equilibrium is characterized by the solution $(x^*(\Delta), y^*(\Delta))$ of the equations

$$u_1(y_1^*(\Delta)) = \delta^{\gamma_1\Delta}u_1(x_1^*(\Delta)) \quad \text{and} \quad u_2(x_2^*(\Delta)) = \delta^{\gamma_2\Delta}u_2(y_2^*(\Delta))$$

(see (3.7)). An argument like that in the proof of Proposition 4.2 shows that the limit, as $\Delta \to 0$, of the agreement $x^*(\Delta)$ is the agreement

$$\arg\max_{(x_1,x_2)\in X} [u_1(x_1)]^{\alpha}[u_2(x_2)]^{1-\alpha},$$

where $\alpha = \gamma_2/(\gamma_1 + \gamma_2)$.

Once again the outcome is given by an asymmetric Nash solution; in this case the exponents reflect a difference in the real time that passes between a rejection and a counteroffer by each player, rather than a difference in the way the players value that time. Notice that the outcome favors the player who can make a counteroffer more quickly. In the extreme case in which $\gamma_i = 0$ the outcome of bargaining is the same as that of the model in which only Player i makes offers.

4.5 A Model with Both Time Preference and Risk of Breakdown

Here we briefly consider a model that combines those in Sections 4.2 and 4.4. In any period, if a player rejects an offer then there is a fixed positive probability that the negotiation terminates in the breakdown event B. The players are not indifferent about the timing of an agreement, or of the breakdown event. Each player's preferences over lotteries on $((X \cup \{B\}) \times T_\infty) \cup \{D\}$ satisfy the assumptions of von Neumann and Morgenstern, and their preferences over this set satisfy C1 through C6. In

addition, for $i = 1, 2$ there is an agreement $b^i \in X$ such that Player i is indifferent between (b^i, t) and (B, t) for all t. Denote by $\Gamma(q, \Delta)$ the game of alternating offers in which the delay between periods is $\Delta > 0$, the breakdown event occurs with probability $q > 0$ after any rejection, and the players' preferences satisfy the assumptions stated above. Then $\Gamma(q, \Delta)$ has a unique subgame perfect equilibrium, which is characterized by the pair of agreements $(x^*(q, \Delta), y^*(q, \Delta))$ that satisfies the following two conditions, where $q \cdot (x, t) \oplus (1 - q) \cdot (y, s)$ denotes the lottery in which (x, t) occurs with probability q and (y, s) occurs with probability $1 - q$:

$$\begin{aligned}(y^*(q, \Delta), 0) &\sim_1 q \cdot (B, 0) \oplus (1 - q) \cdot (x^*(q, \Delta), \Delta)\\ (x^*(q, \Delta), 0) &\sim_2 q \cdot (B, 0) \oplus (1 - q) \cdot (y^*(q, \Delta), \Delta).\end{aligned}$$

We know that under C1 through C6 there exists $0 < \delta < 1$ and concave functions u_i $(i = 1, 2)$ such that Player i's preferences over $X \times T_\infty$ are represented by $\delta^t u_i(x_i)$. However, in general it is *not* possible to choose a representation of this form with the property that its expected value represents i's preferences over lotteries on $X \times T_\infty$. (Suppose, for example, that i's preferences over $X \times T_\infty$ are represented by $\delta^t x_i$. Then in every other representation of the form $\epsilon^t u_i(x_i)$ we have $u_i(x_i) = (x_i)^{(\log \epsilon)/(\log \delta)}$, so that i's preferences over lotteries on $X \times T_\infty$ can be represented in this way only if they display constant relative risk-aversion over X.) If, nevertheless, there exists δ and a function u_i such that Player i's preferences over lotteries on $X \times T_\infty$ *are* represented as the expected value of $\delta^t u_i(x_i)$, then we have

$$u_1(y_1^*(q, \Delta)) = q u_1(B) + (1 - q)\delta^\Delta u_1(x_1^*(q, \Delta)) \tag{4.9}$$

$$u_2(x_2^*(q, \Delta)) = q u_2(B) + (1 - q)\delta^\Delta u_2(y_2^*(q, \Delta)). \tag{4.10}$$

Now consider the limit of the subgame perfect equilibrium as the length Δ of each period converges to zero. Assume that $q = \lambda\Delta$, so that the probability of breakdown in any given interval of real time remains constant. We can then rewrite (4.9) and (4.10) as

$$\begin{aligned}u_1(y_1^*(\Delta)) - \kappa(\Delta)u_1(B) &= \delta^\Delta(1 - \lambda\Delta)\left[u_1(x_1^*(\Delta)) - \kappa(\Delta)u_1(B)\right]\\ u_2(x_2^*(\Delta)) - \kappa(\Delta)u_2(B) &= \delta^\Delta(1 - \lambda\Delta)\left[u_2(y_2^*(\Delta)) - \kappa(\Delta)u_2(B)\right],\end{aligned}$$

where $\kappa(\Delta) = \lambda\Delta/[1 - \delta^\Delta(1 - \lambda\Delta)]$. It follows that

$$\begin{aligned}&(u_1(y_1^*(\Delta)) - \kappa(\Delta)u_1(B))\,(u_2(y_2^*(\Delta)) - \kappa(\Delta)u_2(B)) =\\ &\qquad (u_1(x_1^*(\Delta)) - \kappa(\Delta)u_1(B))\,(u_2(x_2^*(\Delta)) - \kappa(\Delta)u_2(B)).\end{aligned}$$

Notice that if the players use strategies that never lead to agreement, then (given that $q > 0$) with probability one the breakdown event oc-

curs in some period (and D occurs with probability zero). Since $\kappa(\Delta) = \sum_{t=0}^{\infty} \delta^{\Delta t} \lambda\Delta(1-\lambda\Delta)^t$, it follows that $\kappa(\Delta)u_i(B)$ is precisely the expected utility of Player i in this case. Now, letting $r = -\log\delta$, so that $\delta^{\Delta} = e^{-r\Delta}$, we have $\lim_{\Delta\to 0} \kappa(\Delta) = \lambda/(\lambda+r)$. An argument like that in Proposition 4.2 shows that $x^*(\Delta)$ and $y^*(\Delta)$ converge to the Nash solution of the bargaining problem in which the disagreement point is $[\lambda/(\lambda+r)]\,(u_1(B), u_2(B))$, and the agreement set is constructed using the utility functions u_i which, in the special case we are considering, reflect both time preferences *and* risk preferences. This result supports our earlier findings: if δ is close to one (r is close to zero), so that the fear of breakdown rather than the time cost of bargaining is the dominant consideration, then the disagreement point is close to $(u_1(B), u_2(B))$, while if λ is close to zero it is close to $(0, 0)$.

4.6 A Guide to Applications

In order to use a bargaining model as a component of an economic model, we need to choose the economic elements that correspond to the primitives of the bargaining model. The results of this chapter can aid our choice.

4.6.1 Uncertainty as the Incentive to Reach an Agreement

Suppose that we have an economic model in which the main force that causes the parties to reach an agreement is the fear that negotiations will break down. In this case the models of Sections 4.2 and 4.3 indicate that we can apply the Nash solution to an appropriately defined bargaining problem $\langle S, d\rangle$. We should use utility functions that represent the players' preferences over lotteries on the set of physical agreements to construct the set S, and let the disagreement point correspond to the event that occurs if the bargaining is terminated exogenously. By contrast, as we saw in Section 3.12, it is definitely *not* appropriate to take as the disagreement point an outside *option* (an outcome that may or may not occur depending on the choice made by one of the parties).

Suppose, for example, that a buyer and seller are negotiating a price. Assume that they face a risk that the seller's good will become worthless. Assume also that the seller has a standing offer (from a third party) to buy the good at a price that is lower than that which she obtains from the buyer when the third party does not exist. In this case we can apply the Nash solution to a bargaining problem in which the disagreement point reflects the parties' utilities in the event that the good is worthless, and *not* their utilities in the event that the seller chooses to trade with the third party.

4.6.2 Impatience as the Incentive to Reach an Agreement

If the main pressure to reach an agreement is simply the players' impatience, then the original bargaining game of alternating offers studied in Chapter 3 is appropriate. If each player's preferences have the property that the loss to delay is concave (in addition to satisfying all the conditions of Chapter 3), then the result of Section 4.4 shows how the formula for the Nash solution can be used to calculate the limit of the agreement reached in the subgame perfect equilibrium of a bargaining game of alternating offers as the period of delay converges to zero. In this case the utility functions used to construct the set S are concave functions u_i with the property that $\delta^t u_i(x_i)$ represents Player i's preferences ($i = 1, 2$) for some value of $0 < \delta < 1$. Player i's disagreement utility of zero is his utility for an agreement with respect to the timing of which he is indifferent (see C3). Three points are significant here. First, the utility functions of the players are *not* the utility functions they use to evaluate uncertain prospects. Second, if we represent the players' preferences by $\delta_1^t w_1(x_1)$ and $\delta_2^t w_2(x_2)$, where $\delta_1 \neq \delta_2$, and construct the set S using the utility functions w_1 and w_2, then the limit of the agreement reached is given by an *asymmetric* Nash solution in which the exponents depend only on δ_1 and δ_2. Third, the disagreement point does not correspond to an outcome that may occur if the players fail to agree; rather it is determined by their time preferences.

As an example, consider bargaining between a firm and a union. In this case it may be that the losses to the delay of an agreement are significant, while the possibility that one of the parties will find another partner can be ignored. Then we should construct S as discussed above; the disagreement point should correspond to an outcome H with the property that each side is indifferent to the period in which H is received. It might be appropriate, for example, to let H be the outcome in which the profit of the firm is zero and the union members receive a wage that they regard as equivalent to the compensation they get during a strike.

Notes

The basic research program studied in this chapter is the "Nash program" suggested by Nash (1953). When applied to bargaining, the Nash program calls for "supporting" an axiomatic solution by an explicit strategic model of the bargaining process.

Binmore was the first to observe the close relationship between the subgame perfect equilibrium outcome of a bargaining game of alternating offers and the Nash solution (see Binmore (1987a)). The delicacy of the analysis with respect to the distinction between the preferences over lotteries un-

derlying the Nash solution and the time preferences used in the model of alternating offers is explored by Binmore, Rubinstein, and Wolinsky (1986). Our analysis in Sections 4.2, 4.4, and 4.6 follows that paper.

The Demand Game discussed in Section 4.3 is proposed by Nash (1953), who outlines an argument for the result proved there. His analysis is clarified by Binmore (1987a, 1987c) and by van Damme (1987).

Roth (1989) further discusses the relationship between the subgame perfect equilibrium of the game with breakdown and the Nash solution, and Herrero (1989) generalizes the analysis of this relationship to cases in which the set of utilities is not convex. McLennan (1988) generalizes the analysis by allowing nonstationary preferences. Carlsson (1991) studies a variation of the perturbed demand game studied in Section 4.3.

Other games that implement axiomatic bargaining solutions are studied by Howard (1992) (the Nash solution), Moulin (1984) (the Kalai–Smorodinsky solution) and Dasgupta and Maskin (1989) and Anbarci (1993) (the solution that selects the Pareto efficient point on the line through the disagreement point that divides the set of individually rational utility pairs into two equal areas). (Howard's game is based closely on the ordinal characterization of the Nash bargaining solution discussed at the end of Section 2.3.)

CHAPTER 5

A Strategic Model of Bargaining between Incompletely Informed Players

5.1 Introduction

A standard interpretation of the bargaining game of alternating offers studied in Chapter 3 involves the assumption that all players are completely informed about all aspects of the game. In this chapter we modify the model by assuming that one player is completely informed about all aspects of the game, while the other is unsure of the preferences of his opponent.

When each player has complete information about his opponent's preferences, it is not implausible that agreement will be reached immediately. When information is incomplete, however, this is no longer so. Indeed, one of the main reasons for studying models of bargaining between incompletely informed players is to explain delays in reaching an agreement.

When the players in a bargaining game of alternating offers are incompletely informed, they may use their moves as messages to communicate with each other. Each player may try to deduce from his opponent's moves the private information that the opponent possesses; at the same time, he may try to make his opponent believe that he is in a better bargaining

position than he really is. Thus in the analysis of such a model, the issues studied in the literature on signaling come to the forefront.

As in Chapter 3, we formulate the model of bargaining as an extensive game. Following Harsanyi (1967), we convert a situation in which the players are incompletely informed into a game with imperfect information. The fact that information is imperfect means that the notion of subgame perfect equilibrium has little power. For this reason, we appeal to the stronger notion of sequential equilibrium, due to Kreps and Wilson (1982). However, as we shall see in Section 5.3, the set of sequential equilibria is enormously large. In Section 5.4 we study the set and find that it contains outcomes in which agreement is reached only after significant delay. In Section 5.5 we refine the notion of sequential equilibrium by imposing restrictions on the beliefs that the players may entertain when "unexpected" events occur. This refinement gives us a more informative result. However, this result does not accomplish the goal of explaining delay: in any sequential equilibrium satisfying the restrictions on beliefs, there is no significant delay before an agreement is reached. Finally, in Section 5.6 we relate the strategic approach to bargaining between incompletely informed players to the approach taken by the literature on "mechanism design".

5.2 A Bargaining Game of Alternating Offers

The basic model of this chapter is closely related to that of Chapter 3. Two players bargain over the division of a "pie" of size 1. The set of possible agreements is

$$X = \{(x_1, x_2) \in \mathcal{R}^2 \colon x_1 + x_2 = 1 \text{ and } x_i \geq 0 \text{ for } i = 1, 2\}.$$

The players alternately propose agreements at times in $T = \{0, 1, \dots\}$, exactly as in the model of Chapter 3. If the agreement x is accepted in period t, then the outcome is (x, t). The outcome in which an agreement is never reached is denoted D. We restrict attention to the case in which each player has time preferences with a constant cost of delay (see Section 3.3.3). Specifically, Player i's preferences over $X \times T$ are represented by the utility function $x_i - c_i t$ for $i = 1$, 2, and the utility of the disagreement outcome D is $-\infty$. We refer to c_i as Player i's *bargaining cost*.

The basic model departs from that of Chapter 3 in assuming that Player 1 is uncertain of Player 2's bargaining cost. This cost c_2 may take one of the two values c_L and c_H, where $0 < c_L < c_1 < c_H$. We assume that the costs of bargaining are small enough that $c_1 + c_L + c_H < 1$. With probability π_H, Player 2's bargaining cost is c_H, and with probability $1 - \pi_H$ it is c_L. We assume that $0 < \pi_H < 1$. Player 2 knows his own bargaining cost, as well as that of Player 1.

Our assumption that $c_L < c_1 < c_H$ means that Player 1 is in a weak position when matched with an opponent with bargaining cost c_L and in a strong position when matched with an opponent with bargaining cost c_H. In fact, recall that when the players' preferences have fixed bargaining costs, the outcome of the unique subgame perfect equilibrium when the players are completely informed is extreme. When all the bargaining costs are relatively small and it is common knowledge that Player 2 has bargaining cost c_L, Player 1 obtains a small payoff; it is positive only because Player 1 has the advantage of being the first to make a proposal. If it is common knowledge that Player 2's bargaining cost is c_H then Player 1 obtains all the pie (see Section 3.9.2). Thus in the game in which Player 1 is unsure of Player 2's type, Player 2 has every incentive to convince Player 1 that his bargaining cost is c_L.

We represent this situation as an extensive game by introducing *two* players in the role of Player 2. One of these, whom we call 2_L, has bargaining cost c_L, while the other, whom we call 2_H, has bargaining cost c_H. Player 1 does not know which of these players she faces. At the beginning of the game, Player 2_H is selected with probability π_H, and Player 2_L is selected with probability $1-\pi_H$. Given the outcomes in the games of complete information between Players 1 and 2_H, and between Players 1 and 2_L, we refer to Player 2_H as "weak" and to Player 2_L as "strong". Following convention we sometimes refer to 2_H and 2_L as *types* of Player 2.

A representation of the game, which we denote $\Gamma(\pi_H)$, is shown in Figure 5.1. The fact that Player 1 is not informed of the selection of Player 2's bargaining cost is indicated by the dotted line connecting the two decision nodes of Player 1 at $t = 0$.[1] The first decision in the game is Player 1's; she proposes an agreement to Player 2. In Figure 5.1 one such proposal x^0 is indicated. Subsequently, each of Players 2_H and 2_L either accept or reject the proposal. If it is accepted, then the game ends with the outcome $(x^0, 0)$. If it is rejected, then play moves to the next period, in which Player 2 makes a counteroffer, which may depend on his type and on Player 1's rejected proposal. Player 1 observes this counteroffer but cannot tell whether it came from Player 2_H or Player 2_L. Thus Player 1's response at $t = 1$ may depend upon both the offer at $t = 1$ and the rejected offer at $t = 0$. The case in which these offers are x^1 and x^0 is shown in Figure 5.1. If Player 1 accepts the counteroffer, then the game ends; if she rejects it, then play passes to period 2, in which it is again her turn to propose an agreement.

A *history* is a sequence of proposals and responses. A *strategy* of each player in $\Gamma(\pi_H)$ specifies an action for every possible history after which he

[1]In the language of game theory, the two nodes constitute an *information set* of Player 1.

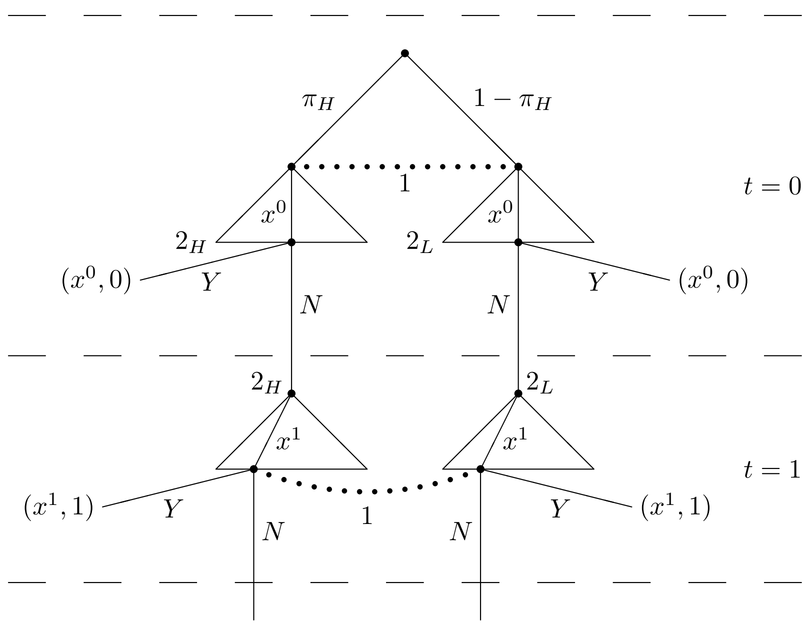


Figure 5.1 The first two periods of the game $\Gamma(\pi_H)$. The game begins by the selection of Player 2's bargaining cost. The fact that Player 1 is not informed of this selection is indicated by the dotted line that connects the first two nodes at which Player 1 has to make a choice. The branches labeled x^0 represent a "typical" offer of Player 1 out of the continuum available in period 0; similarly the branches labeled x^1 represent a "typical" offer of Player 2 in period 1.

has to move.[2] Thus a strategy of Player 1 has exactly the same structure as a strategy of Player 1 in the game studied in Chapter 3 (see Section 3.4), and strategies for Players 2_L and 2_H each have precisely the same form as a strategy for Player 2 in that game.

A triple of strategies, one each for Players 1, 2_H, and 2_L, leads, from the point of view of Player 1, to a probability distribution over outcomes. With probability π_H the outcome is that given by the combination of the strategies of Players 1 and 2_H, while with probability $1 - \pi_H$ the outcome is that given by the combination of the strategies of Players 1 and 2_L. In order to compare two of her strategies, Player 1 thus has to compare two probability distributions over outcomes. Hence we must extend the domain

[2] As in Chapter 3 we do not allow players to use a random device to select their actions.

of her preferences from $(X \times T) \cup \{D\}$ to lotteries over this space. We do so by assuming that Player 1 evaluates each lottery by its expected utility.

5.3 Sequential Equilibrium

The notion of Nash equilibrium (see Section 3.6) can be applied in a straightforward manner to the game $\Gamma(\pi_H)$. However, as in the game studied in Chapter 3, in which each player is fully informed, $\Gamma(\pi_H)$ has a great multiplicity of Nash equilibria. In Chapter 3, we isolated a unique solution by requiring that each player's action be optimal from any point on, not just at the start of the game. For games in which the players are imperfectly informed, this idea is embodied in the notion of sequential equilibrium.[3]

In order to state the requirement that Player 1's strategy be optimal for every history after which she has to move we must specify her probabilistic beliefs about Player 2's type. (Notice that Player 2's type is the only element of uncertainty for Player 1). Therefore, the notion of sequential equilibrium requires us to specify *two* elements: the profile of strategies *and* the beliefs of Player 1.

A *system of beliefs* in $\Gamma(\pi_H)$ is a function p_H that assigns a number $p_H(h) \in [0, 1]$ to every history h after which Player 1 has to move. The number $p_H(h)$ is interpreted as the probability that Player 1 assigns, after the history h, to the event that her opponent is 2_H.

We impose three conditions on the pair of strategies and beliefs. First, we require that each player's strategy be optimal after every history. We refer to this condition as "sequential rationality". The optimality of Player 1's strategy after any history h depends on the strategies of Players 2_H and 2_L and on her beliefs after h. Since Player 2 is perfectly informed, the optimality of the strategies of Players 2_H and 2_L after any history h depends only on Player 1's strategy.

The second condition, which we refer to as "consistency", is closely related to the consistency condition of Kreps and Wilson (1982).[4] It requires that Player 1's beliefs be consistent with the probability π_H with which she initially faces Player 2_H and with the strategies of Players 2_H and 2_L. As play proceeds, Player 1 must, whenever possible, use Bayes' rule to update her beliefs. If, after any history, the strategies of Players 2_H and 2_L call for them both to reject an offer and make the same counteroffer, and this counteroffer is indeed made, then when responding to the counteroffer Player 1's belief remains the same as it was when she made the offer. If only

[3]The notion of subgame perfect equilibrium which we defined in Chapter 3 has no power in $\Gamma(\pi_H)$, since this game has no proper subgames.

[4]For a discussion of the condition see Kreps and Ramey (1987).

one of the strategies of Players 2_H and 2_L specifies that the offer made by Player 1 be rejected and the counteroffer x be made, and the counteroffer x is indeed made, then when responding to the counteroffer Player 1's belief is zero or one, as appropriate. If neither of the strategies of Players 2_H and 2_L call for them to reject an offer and make some counteroffer x, but the counteroffer x is observed, then Player 1 cannot use Bayes' rule to update her belief. In this case she may choose any number in the interval $[0, 1]$ as her belief. In all cases Player 1's belief when she makes a counteroffer (after rejecting the counteroffer of Player 2) must be the same as it was when she responded to Player 2's counteroffer: only actions of Player 2 lead Player 1 to change her belief. To summarize, after any history Player 1's beliefs must be based on her previous beliefs and Player 2's strategies as long as the response and counteroffer of Player 2 are consistent with the strategy of either Player 2_H or Player 2_L, or both. Whenever Player 2's response or counteroffer is inconsistent with both of these strategies, Player 1 is free to form new beliefs.

Three points are worth noting about this condition. First, having formed a new belief after an action of Player 2 that is inconsistent with the strategies of both Player 2_H and Player 2_L, Player 1 is required subsequently to update this belief in accordance with the strategies of Players 2_H and 2_L. This is possible since the strategy of each player specifies his behavior after he takes an action inconsistent with the strategy. Second, Player 1 updates her beliefs *only* when she is about to take an action. If, for example, the strategy of only one of the types of Player 2 calls for him to reject an offer of Player 1, and Player 1's offer is indeed rejected, then Player 1 does *not* necessarily conclude that she faces that type unless she also receives the counteroffer prescribed by that type's strategy. Third, the condition implicitly requires that a deviation by Player 1 herself not affect the belief she uses as the basis for her updating. Thus, for example, if she proposes an agreement in period 0 different from that specified by her strategy, she must still use the initial probability π_H as the basis for her updating when she responds to Player 2's counteroffer.

The last condition we impose is the following. Once Player 1 is convinced of the identity of Player 2, she is never dissuaded from her view. We refer to this condition as NDOC ("Never Dissuaded Once Convinced"). The condition implies, for example, that once Player 1 reaches the conclusion that she faces Player 2_L with certainty ($p_H(h) = 0$), she cannot revise her belief, even if Player 2 subsequently deviates from the strategy of Player 2_L: from this point on she is engaged in a game of perfect information with Player 2_L.[5]

[5]For a discussion of this constraint see Madrigal, Tan, and Werlang (1987).

NDOC is a strong assumption. Sometimes circumstances lead one to retreat from a belief that with certainty one faces a given type of player. However, if we allow a player in a game of incomplete information to change his mind after he has been persuaded that he is playing with certainty against a given type, then why we do not do so in a game of complete information? The issue is unclear; more research is needed to clarify it.

To summarize, we make the following definition.

Definition 5.1 A *sequential equilibrium* of $\Gamma(\pi_H)$ is a pair consisting of a triple of strategies (one each for Players 1, 2_L, and 2_H) and a system of beliefs p_H that satisfies the following properties.

Sequential Rationality After every history for which Player 1 has to move, her strategy is optimal, given the strategies of Players 2_H and 2_L, and given p_H. After every history for which Player 2 has to move, Player 2_I's strategy ($I = H, L$) is optimal (with respect to his preferences), given Player 1's strategy.

Consistency The initial belief is π_H. Let $h = (x^0, N, x^1, N, \ldots, x^T, N)$, where T is odd, and let $h' = (x^0, N, x^1, N, \ldots, x^{T+1}, N, x^{T+2})$. If, after the history h, the strategies of Players 2_H and 2_L both call for them to reject x^{T+1} and to counteroffer x^{T+2}, then $p_H(h') = p_H(h)$. If $p_H(h) \neq 0$ and only the strategy of 2_H rejects x^{T+1} and counteroffers x^{T+2} then $p_H(h') = 1$; if $p_H(h) \neq 1$ and only the strategy of 2_L rejects x^{T+1} and counteroffers x^{T+2} then $p_H(h') = 0$. Further, $p_H(x^0, N, x^1, N, \ldots, x^{T+2}, N) = p_H(x^0, N, x^1, N, \ldots, x^{T+2})$.

Never Dissuaded Once Convinced (NDOC) If $p_H(h) = 0$ for some history h then it remains zero for all subsequent histories, and if $p_H(h) = 1$ for some history h then it remains one for all subsequent histories.

We now establish some properties of all sequential equilibria of $\Gamma(\pi_H)$.

Lemma 5.2 *After any history h for which $0 < p_H(h) < 1$, every sequential equilibrium of $\Gamma(\pi_H)$ has the following properties.*

1. *If the strategies of Players 2_H and 2_L call for them both to reject an offer, then these strategies also call for them to make the same counteroffer.*
2. *If the strategy of Player 2_L calls for him to accept an offer, so does the strategy of Player 2_H.*
3. *If the strategy of Player 2_H calls for him to accept the proposal x while the strategy of Player 2_L calls for him to reject it, then Player 1's strategy calls for her to accept the counteroffer y that Player 2_L's strategy prescribes, and $x_1 - c_H \le y_1 \le x_1 - c_L$.*

Proof. We prove each part in turn.

1. Assume that for some history the strategies of Players 2_H and 2_L call for them to reject the agreement proposed by Player 1 and make the different counteroffers y and z, respectively. Then the consistency condition demands that if Player 1 is offered y then she believes that she faces Player 2_H with probability one, and if she is offered z then she believes that she faces Player 2_L with probability one. Under condition NDOC this belief never changes subsequently. Thus if Player 1 rejects y, then agreement is reached immediately on $(1, 0)$ (the outcome in the game of perfect information between Players 1 and 2_H), whereas if she rejects z, then agreement is reached immediately on $(c_L, 1 - c_L)$ (the outcome in the game of perfect information between Players 1 and 2_L). Since $c_L - c_1 < 0 \leq z_1$, Player 1 accepts z. If also she accepts y, then one of the types can profitably deviate by proposing either y or z, whichever has the higher share for Player 2. Thus Player 1 must reject y. But then Player 2_H receives 0 with a delay of one period. This is worse than receiving z_2 immediately, which is possible if he imitates Player 2_L and proposes z. Thus it is not optimal for Player 2_H to offer y, contradicting our original assumption.

2. Suppose that the strategy of Player 2_L calls for him to accept a proposal, while that of Player 2_H calls for him to reject the same proposal. Then it is better for Player 2_H to deviate and accept the proposal of Player 1 since by doing so he obtains at least 0, while if he follows his strategy and rejects the offer then Player 1 concludes that she faces Player 2_H, so that (under the condition NDOC) the outcome is immediate agreement on $(1 - c_1, c_1)$, which yields Player 2_H a payoff of $c_1 - c_H < 0$.

3. If Player 2 rejects the offer x and proposes y, then Player 1 concludes that she faces Player 2_L, so that it is optimal for her to accept y. If Player 2_H deviates and imitates Player 2_L, then he obtains y_2 with one period of delay, instead of the x_2 he gets when he accepts Player 1's offer. Thus we must have $x_2 \geq y_2 - c_H$, or $y_1 \geq x_1 - c_H$. Similarly, we must have $y_2 - c_L \geq x_2$, or $y_1 \leq x_1 - c_L$, in order for it not to be profitable for Player 2_L to imitate Player 2_H and accept x. □

As we noted above, in the game $\Gamma(\pi_H)$ the notion of sequential equilibrium puts no restriction on Player 1's belief about the opponent she faces when an unexpected event occurs. The next result shows that unless π_H is high, this freedom to specify beliefs leads to a great multiplicity of equilibria.

Proposition 5.3

1. *If* $\pi_H > 2c_1/(c_1+c_H)$ *then the minimum of Player 1's expected payoff over all sequential equilibria of* $\Gamma(\pi_H)$ *is* $\pi_H + (1-\pi_H)(1-c_H-c_1)$.
2. *If* $\pi_H \leq 2c_1/(c_1+c_H)$ *then for every* $\xi^* \in [c_1, 1-c_1+c_L]$ *there is a ("pooling") sequential equilibrium of* $\Gamma(\pi_H)$ *in which Player 1 proposes* $x^* = (\xi^*, 1-\xi^*)$ *in period 0, which Players* 2_H *and* 2_L *both immediately accept.*
3. *If* $(c_1+c_L)/(c_1+c_H) \leq \pi_H \leq 2c_1/(c_1+c_H)$ *then for every* $\xi^* \geq c_H$ *there is a ("separating") sequential equilibrium of* $\Gamma(\pi_H)$ *in which Player 1 proposes* $x^* = (\xi^*, 1-\xi^*)$ *in period 0, Player* 2_H *accepts* x^*, *and Player* 2_L *rejects it and proposes* $(\xi^* - c_H, 1-\xi^*+c_H)$, *which Player 1 accepts.*

Part 1 of the proposition shows that if it is likely that Player 2 is weak (i.e. has the high bargaining cost) then, when the bargaining costs are small, Player 1 gets a large share of the pie in all sequential equilibria of $\Gamma(\pi_H)$.[6] Parts 2 and 3 show that when the probability that Player 2 is weak is relatively small, however, the notion of sequential equilibrium is very uninformative: almost every agreement to be the outcome of a sequential equilibrium. For example, even if π_H is close to zero there is a sequential equilibrium in which, when the bargaining costs are small, Player 1 obtains almost all the pie. In the equilibria we construct to establish the result, Player 1 believes, after any deviation, that she faces the weak player (and acts accordingly). More precisely, whenever an event occurs that is inconsistent with the equilibrium strategies of both Player 2_H and Player 2_L, Player 1 makes the "optimistic" conjecture that her opponent is Player 2_H with probability one. This optimistic conjecture gives credibility to a tough bargaining strategy for Player 1 and allows a wide range of equilibria to be generated: if Player 2 deviates then the switch in Player 1's belief leads her to persistently demand the whole pie, which deters the deviation.

Proof of Proposition 5.3. We proceed in steps.

Step 1. The strategies and beliefs described in Table 5.1 constitute a sequential equilibrium of $\Gamma(\pi_H)$ for $(c_1+c_L)/(c_1+c_H) \leq \pi_H \leq 2c_1/(c_1+c_H)$ if $x_1^* \geq c_H$, and for $\pi_H \geq (c_1+c_L)/(c_1+c_H)$ if $x_1^* = 1$.

Proof. Note that in state I, for $I = H$, L, Players 1 and 2_I behave as they do in the unique subgame perfect equilibrium of the complete information game between Player 1 and Player 2_I, and the other type of Player 2

[6] This part corrects a mistake in Part 2 of Proposition 4 of Rubinstein (1985b).

		x^*	H	L
1	proposes	x^*	$(1, 0)$	$(c_L, 1 - c_L)$
	accepts	$x_1 \geq x_1^* - c_H$	$x_1 \geq 1 - c_1$	$x_1 \geq 0$
	belief	π_H	1	0
2_H	proposes	$(x_1^* - c_H, x_2^* + c_H)$	$(1 - c_1, c_1)$	$(0, 1)$
	accepts	$x_1 \leq x_1^*$	$x_1 \leq 1$	$x_1 \leq c_H$
2_L	proposes	$(x_1^* - c_H, x_2^* + c_H)$	$(1 - c_1, c_1)$	$(0, 1)$
	accepts	$x_1 \leq x_1^* - c_H + c_L$	$x_1 \leq 1 - c_1 + c_L$	$x_1 \leq c_L$
	Transitions	Go to L if Player 2 rejects x with $x_1^* - c_H + c_L < x_1 \leq x_1^*$ and counterproposes $(x_1^* - c_H, x_2^* + c_H)$.	Absorbing	Absorbing
		Go to H if Player 2 takes an action inconsistent with the strategies of both 2_H and 2_L.		

Table 5.1 A ("separating") sequential equilibrium of $\Gamma(\pi_H)$ for $(c_1 + c_L)/(c_1 + c_H) \leq \pi_H \leq 2c_1/(c_1 + c_H)$. The value of x^* satisfies $c_H \leq x_1^* \leq 1$. When $x_1^* = 1$ the strategy profile is a sequential equilibrium also for $\pi_H > 2c_1/(c_1 + c_H)$.

uses a best response to the strategy of Player 1. Following our convention, the initial state is the one in the leftmost column, namely x^*. As always, transitions between states occur immediately after the event that triggers them. Thus the transition to state L occurs after Player 2 makes an offer, *before* Player 1 responds, and, for example, a response of Player 2 that is inconsistent with the strategies of both Player 2_H and Player 2_L causes a transition to state H *before* Player 2 makes a counteroffer. (Refer to Section 3.5 for a discussion of this method of representing an equilibrium.[7] An

[7]The representation of the strategies presented in Table 5.1 as standard automata is more complex than the representation for the example given in Section 3.5, since the transition to state L depends on both the counterproposal *and* the previously rejected offer. For each state in the table, we need to introduce a set of states indexed by i and x in which Player i has to respond to the offer x, and another set indexed by the same variables in which Player i has to make a counteroffer, given that the previously rejected proposal was x.

extra line is included for Player 1, since the notion of sequential equilibrium includes a specification of Player 1's belief as well as her actions.)

To see that players' behavior in state x^* is optimal, first consider Player 1. The best proposal out of those that are accepted by both Player 2_H and Player 2_L is that in which $x_1 = x_1^* - c_H + c_L$. This results in a payoff for Player 1 of $x_1^* - c_H + c_L$, which, under our assumption that $\pi_H \geq (c_1 + c_L)/(c_1 + c_H)$, is at most equal to Player 1's equilibrium payoff of $\pi_H x_1^* + (1 - \pi_H)(x_1^* - c_H - c_1)$. If Player 1 proposes an agreement x in which $x_1 > x_1^*$, then this proposal is rejected by both Player 2_H and Player 2_L, who counterpropose $(x_1^* - c_H, x_2^* + c_H)$, which Player 1 accepts, yielding her a payoff of $x_1^* - c_H - c_1$. If $x_1^* = 1$ then Player 1's acceptance rule in state x^* is never activated: after any counteroffer of Player 2 in period 1, there is a transition to either state L or state H. If $x_1^* < 1$ then the only offer that Player 1 is confronted with in state x^* gives her $x_1^* - c_H$. If she rejects this offer then she counterproposes x^* and obtains her equilibrium payoff with one period of delay, the value of which is at most $x_1^* - c_H$ if $\pi_H \leq 2c_1/(c_1 + c_H)$.

Now consider the behavior of Player 2_L. If in state x^* he rejects an offer x in which $x_1 > x_1^* - c_H + c_L$, then he counterproposes $(x_1^* - c_H, x_2^* + c_H)$, which Player 1 accepts. (If $x_1 \geq x_1^*$ then the state changes to L before Player 1's acceptance.) Thus it is optimal to reject such an offer. If he rejects an offer x in which $x_1 \leq x_1^* - c_H + c_L$, then the state changes to H, and he obtains $c_1 - c_L < x_2^* + c_H - c_L$, so it is optimal to accept. Now consider his proposal in state x^*. Let the offer he rejected previously be x. We must have $x_1 > x_1^* - c_H + c_L$, otherwise there would have been a transition to state H. Thus if he proposes $(x_1^* - c_H, x_2^* + c_H)$ then if $x_1 \leq x_1^*$ the state changes to L, while if $x_1 > x_1^*$ the state remains x^*; in both cases Player 1 will accept the offer. If he proposes y with $y_1 \neq x_1^* - c_H$, then the state changes to H. If $y_1 < 1 - c_1$, then Player 1 rejects the offer, and Player 2_L obtains $c_1 - 2c_L$; if $y_1 \geq 1 - c_1$, then Player 1 accepts the offer, and Player 2_L obtains at most c_1. Thus in both cases it is better to propose $(x_1^* - c_H, x_2^* + c_H)$. (Note that Player 1 does *not* conclude, after Player 2_L rejects an offer x with $x_1^* - c_H + c_L < x_1 \leq x_1^*$, that she faces Player 2_L. She is required by the consistency condition to draw this conclusion *only* after Player 2_L makes the counteroffer $(x_1^* - c_H, x_2^* + c_H)$.)

The optimality of Player 2_H's strategy in state x^*, and of the strategies in the other states can be checked similarly. Finally, the postulated beliefs are consistent with the strategies.

This completes the proof of Part 3 of the proposition.

Now let m_1 be the infimum of Player 1's payoffs in all sequential equilibria of the game $\Gamma_1(\pi_H)$ (which is the same as $\Gamma(\pi_H)$) starting with an offer by

Player 1 in which the initial belief of Player 1 is π_H, and let M_H be the supremum of Player 2_H's payoffs in all sequential equilibria of the game $\Gamma_2(\pi_H)$ starting with an offer by Player 2 in which the initial belief of Player 1 is π_H. The first two steps follow the lines of Steps 1 and 2 in the proof of Theorem 3.4.

Step 2. $m_1 \geq \pi_H(1-\max\{M_H-c_H, 0\})+(1-\pi_H)(1-\max\{M_H - c_H, 0\}-c_H - c_1)$.

Proof. Suppose that in $\Gamma_1(\pi_H)$, Player 1 proposes an agreement x in which $x_2 > \max\{M_H-c_H, 0\}$. If Player 2_H rejects x, then so does Player 2_L (by Part 2 of Lemma 5.2), and both types make the same counteroffer (by Part 1 of Lemma 5.2), so that play passes to the game $\Gamma_2(\pi_H)$. In this game Player 2_H receives at most M_H, so that in any sequential equilibrium he must accept x. If Player 2_L rejects x, then by Part 3 of Lemma 5.2, he proposes an agreement y with $y_1 \geq x_1 - c_H$, which Player 1 accepts. Thus by proposing the agreement x with x_2 sufficiently close to $\max\{M_H-c_H, 0\}$, Player 1 can obtain a payoff arbitrarily close to the amount on the right-hand side of the inequality to be established.

Step 3. $M_H \leq 1 - (m_1 - c_1)$.

Proof. By Part 1 of Lemma 5.2, Players 2_H and 2_L make the same offer in period 0 of $\Gamma_2(\pi_H)$, so that if Player 1 rejects a proposal in period 0, her belief remains π_H, and play passes into the game $\Gamma_1(\pi_H)$, in which Player 1's expected payoff is at least m_1. Thus Player 1's expected payoff in all sequential equilibria of $\Gamma_2(\pi_H)$ is at least $m_1 - c_1$. The inequality we need to establish follows from the fact that in no sequential equilibrium of $\Gamma_2(\pi_H)$ is Player 2_H's payoff higher than that of Player 2_L (since Player 2_L can imitate Player 2_H, and has a lower bargaining cost).

Step 4. If $\pi_H > 2c_1/(c_H + c_1)$ then $m_1 = \pi_H + (1 - \pi_H)(1 - c_H - c_1)$.

Proof. If $M_H > c_H$ then Steps 2 and 3 imply that $1 - M_H + \pi_H(c_H + c_1) - c_1 \leq m_1 \leq 1 - M_H + c_1$, which violates the assumption that $\pi_H > 2c_1/(c_H + c_1)$. Thus $M_H \leq c_H$, so that from Step 2 we have $m_1 \geq \pi_H + (1 - \pi_H)(1 - c_H - c_1)$. Finally, Step 1 (for the case $x_1^* = 1$) shows that $m_1 \leq \pi_H + (1 - \pi_H)(1 - c_H - c_1)$ if $\pi_H \geq (c_1 + c_L)/(c_1 + c_H)$, and hence certainly if $\pi_H > 2c_1/(c_H + c_1)$.

This completes the proof of Part 1 of the proposition.

Step 5. If $\pi_H \leq 2c_1/(c_1 + c_H)$ then the strategies and beliefs described in Table 5.2 constitute a sequential equilibrium of $\Gamma(\pi_H)$ whenever $c_1 \leq x_1^* \leq 1 - c_1 + c_L$.

		x^*	H	L
1	proposes	x^*	$(1, 0)$	$(c_L, 1 - c_L)$
	accepts	$x_1 \geq x_1^* - c_1$	$x_1 \geq 1 - c_1$	$x_1 \geq 0$
	belief	π_H	1	0
2_H	proposes	$(x_1^* - c_1, x_2^* + c_1)$	$(1 - c_1, c_1)$	$(0, 1)$
	accepts	$x_1 \leq x_1^* + c_H - c_1$	$x_1 \leq 1$	$x_1 \leq c_H$
2_L	proposes	$(x_1^* - c_1, x_2^* + c_1)$	$(1 - c_1, c_1)$	$(0, 1)$
	accepts	$x_1 \leq x_1^*$	$x_1 \leq 1 - c_1 + c_L$	$x_1 \leq c_L$
Transitions		Go to L if Player 2 rejects x with $x_1^* < x_1 \leq x_1^* + c_H - c_1$ and counterproposes $(x_1^* - c_1, x_2^* + c_1)$.	Absorbing	Absorbing
		Go to H if Player 2 takes an action inconsistent with the strategies of both 2_H and 2_L.		

Table 5.2 A ("pooling") sequential equilibrium of $\Gamma(\pi_H)$ for $\pi_H \leq 2c_1/(c_1 + c_H)$. The value of x^* satisfies $c_1 \leq x_1^* \leq 1 - c_1 + c_L$.

Proof. Note that the states H and L are the same as for the equilibrium constructed in Step 1 above. To see that the strategies and beliefs constitute a sequential equilibrium, first consider Player 1. If she proposes an agreement x in which $x_1^* < x_1 \leq x_1^* + c_H - c_1$. Then Player 2_H accepts x, while Player 2_L rejects it and proposes the agreement in which Player 1 receives $x_1^* - c_1$, the state changes to L, and Player 1 accepts the offer. Thus by deviating in this way, Player 1 can obtain no more than $\pi_H(x_1^* + c_H - c_1) + (1 - \pi_H)(x_1^* - 2c_1) = x_1^* + \pi_H c_H - c_1(2 - \pi_H)$, which is equal to at most x_1^* by our assumption that $\pi_H \leq 2c_1/(c_1 + c_H)$. If Player 1 proposes an agreement x for which $x_1 > x_1^* + c_H - c_1$ then Players 2_H and 2_L both reject it and counterpropose $(x_1^* - c_1, x_2^* + c_1)$, which Player 1 accepts. Thus Player 1 obtains $x_1^* - 2c_1$, which is less than her payoff if she adheres to her strategy. The only offer that Player 1 can be confronted with in state x^* is $(x_1^* - c_1, x_2^* + c_1)$; if she rejects this then she proposes x^*, which both types of Player 2 accept, so that

she obtains $x_1^* - c_1$, the same payoff that she obtains if she accepts the offer.

If Player 2_L rejects an offer x in which $x_1 < x_1^*$, then the state changes to H, so that Player 2_L obtains $c_1 - c_L$. The condition $x_1^* \leq 1 - c_1 + c_L$ ensures that this payoff is no more than x_2. The fact that no player can benefit from any other deviation can be checked similarly. Finally, the postulated beliefs are consistent with the strategies.

This completes the proof of Part 2 of the proposition. □

5.4 Delay in Reaching Agreement

In Chapter 3 we found that in the unique subgame perfect equilibrium of a bargaining game of alternating offers in which the players' preferences are common knowledge, agreement is reached immediately. In the previous section we constructed sequential equilibria for the game $\Gamma(\pi_H)$ in which, when Player 1 faces a strong opponent, agreement is reached with delay, but in these equilibria this delay never exceeds one period. Are there any equilibria in which the negotiation lasts for more than two periods? If so, can the bargaining time remain bounded away from zero when the length of a period of negotiation is arbitrarily small?

In the case that $\pi_H \leq 2c_1/(c_1 + c_H)$ we now construct a sequential equilibrium in which negotiation continues for several periods. Choose three numbers $\xi^* < \eta^* < \zeta^*$ from the interval $[c_1, 1 - c_1 + c_L]$ such that $\zeta^* - \eta^* > c_1 - c_L$ (this is possible if the bargaining costs are small), and let $\bar{t}$ be an even integer. Recall that for each $\alpha \in [c_1, 1 - c_1 + c_L]$ there is a sequential equilibrium in which immediate agreement is reached on $(\alpha, 1 - \alpha)$ (by Part 2 of Proposition 5.3). The players' strategies in the equilibrium we construct are as follows. Through period $\bar{t}$, Player 1 proposes the agreement $(1, 0)$ and rejects every other agreement, and Players 2_H and 2_L each propose the agreement $(0, 1)$ and reject every other agreement; Player 1 retains her original belief that the probability with which she faces Player 2_H is π_H. If period $\bar{t}$ is reached without any of the players having deviated from these strategies, then from period $\bar{t}+1$ on the players use the strategies of a sequential equilibrium that leads to immediate agreement on $y^* = (\eta^*, 1 - \eta^*)$. If in any period $t \leq \bar{t}$ Player 1 proposes an agreement different from $(1, 0)$, then subsequently the players use the strategies of a sequential equilibrium that leads to immediate agreement on $x^* = (\xi^*, 1 - \xi^*)$ in the case that Player 1 is the first to make an offer. If Player 2 proposes an agreement different from $(0, 1)$ in some period $t \leq \bar{t}$ then Player 1 retains the belief that she faces Player 2_H with prob-

ability π_H, and subsequently the players use the strategies of a sequential equilibrium that leads to immediate agreement on $z^* = (\zeta^*, 1-\zeta^*)$.

The outcome of this strategy profile is that no offer is accepted until period $\bar{t}+1$. In this period Player 1 proposes y^*, which Players 2_H and 2_L both accept.

In order for these strategies and beliefs to constitute a sequential equilibrium, the number $\bar{t}$ has to be small enough that none of the players is better off making a less extreme proposal in some period before $\bar{t}$. The best such alternative proposal for Player 1 is x^*, and the best period in which to make this proposal is the first. If she deviates in this way, then she obtains x_1^* rather than $y_1^* - c_1\bar{t}$. Thus we require $\bar{t} \leq (y_1^* - x_1^*)/c_1$ in order for the deviation not to be profitable. The best deviation for Player 2_I ($I = H, L$) is to propose $(z_1^* - c_1, 1 - z_1^* + c_1)$ in the second period (the first in which he has the opportunity to make an offer). In the equilibrium, Player 1 accepts this offer, so that Player 2_I obtains $1 - z_1^* + c_1 - c_I$ rather than $1 - y_1^* - c_I\bar{t}$. Thus in order to prevent a deviation by either Player 2_H or Player 2_L we further require that $\bar{t} \leq (z_1^* - y_1^* + c_I - c_1)/c_I$ for $I = H, L$.

We can interpret the equilibrium as follows. The players regard a deviation as a sign of weakness, which they "punish" by playing according to a sequential equilibrium in which the player who did not deviate is better off. Note that there is delay in this equilibrium even though no information is revealed along the equilibrium path.

Now consider the case in which a period has length Δ. Let Player 1's bargaining cost be $\gamma_1\Delta$ per period, and let Player 2_I's be $\gamma_I\Delta$ for $I = H$, L. Then the strategies and beliefs we have described constitute a sequential equilibrium in which the real length $\bar{t}\Delta$ of the delay before an agreement is reached can certainly be as long as the minimum of $(y_1^* - x_1^*)/\gamma_1$, $(z_1^* - y_1^* + \gamma_H\Delta - \gamma_1\Delta)/\gamma_H$, and $(z_1^* - y_1^* + \gamma_L\Delta - \gamma_1\Delta)/\gamma_L$. The limit of this delay, as $\Delta \to 0$, is positive, and, if the bargaining cost of each player is relatively small, can be long. Thus if $\pi_H < 2c_1/(c_1 + c_H)$, a significant delay is consistent with sequential equilibrium even if the real length of a period of negotiation is arbitrarily small.

In the equilibrium we have constructed, Players 2_H and 2_L change their behavior after a deviation and after period $\bar{t}$ is reached, even though Player 1's beliefs do not change. Gul and Sonnenschein (1988) impose a restriction on strategies that rules this out. They argue that the offers and response rules given by the strategies of Players 2_H and 2_L should depend only on the belief held by Player 1, and not, for example, on the period. We show that among the set of sequential equilibria in which the players use strategies of this type, there is no significant delay before an agreement is reached.

Proposition 5.4 *In any sequential equilibrium in which the offers and*

response rules given by the strategies of Players 2_H and 2_L depend only on the belief of Player 1, agreement is reached not later than the second period.

Proof. Since the cost of perpetual disagreement is infinite, all sequential equilibria must end with an agreement. Consider a sequential equilibrium in which an agreement is first accepted in period $\bar{t} \geq 2$. Until this acceptance, it follows from Part 1 of Lemma 5.2 that in any given period t, Players 2_H and 2_L propose the same agreement y^t, so that Player 1 continues to maintain her initial belief π_H. Hence, under the restriction on strategies, the agreement y^t, and the acceptance rules used by Players 2_H and 2_L, are independent of t. Thus if it is Player 1 who first accepts an offer, she is better off deviating and accepting this offer in the second period, rather than waiting until period $\bar{t}$. By Lemma 5.2 the only other possibility is that Player 2_H accepts x in period $\bar{t}$ and Player 2_L either does the same, or rejects x and makes a counterproposal that is accepted. By the restriction on the strategies Player 2_L's counterproposal is independent of $\bar{t}$. Thus in either case Player 1 is better off proposing x in period 0. Hence we must have $\bar{t} \leq 1$. □

Gul and Sonnenschein actually establish a similar result in the context of a more complicated model. Their result, as well as that of Gul, Sonnenschein, and Wilson (1986), is associated with the "Coase conjecture". The players in their model are a seller and a buyer. The seller is incompletely informed about the buyer's reservation value, and her initial probability distribution F over the buyer's reservation value is continuous and has support $[l, h]$. Gul and Sonnenschein assume that (i) the buyer's actions depend only on the seller's belief, (ii) the seller's offer after histories in which she believes that the distribution of the buyer's reservation value is the conditional distribution of F on some set $[l, h']$ is increasing in h', and (iii) the seller's beliefs do not change in any period in which the negotiation does not end if all buyers follow their equilibrium strategies. They show that for all $\epsilon > 0$ there exists Δ^* small enough such that in any sequential equilibrium of the game in which the length of a period is less than Δ^* the probability that bargaining continues after time ϵ is at most ϵ.

Gul and Sonnenschein argue that their result demonstrates the shortcomings of the model as an explanation of delay in bargaining. However, note that their result depends heavily on the assumption that the actions of the informed player depend only on the belief of the uninformed player. (This issue is discussed in detail by Ausubel and Deneckere (1989a).) This assumption is problematic. As we discussed in Section 3.4, we view a player's strategy as more than simply a plan of action. The buyer's strategy also includes the seller's predictions about the buyer's behavior in case that the

buyer does not follow his strategy. Therefore the assumption of Gul and Sonnenschein implies not only that the buyer's plan of action is the same after any history in which the seller's beliefs are the same. It implies also that the seller does not make any inference about the buyer's future plans from a deviation from his strategy, unless the deviation also changes the seller's beliefs about the buyer's reservation value.

5.5 A Refinement of Sequential Equilibrium

Proposition 5.3 shows that the set of sequential equilibria of the game $\Gamma(\pi_H)$ is very large. In this section we strengthen the notion of sequential equilibrium by constraining the beliefs that the players are allowed to entertain when unexpected events occur.

To motivate the restrictions we impose on beliefs, suppose that Player 2 rejects the proposal x and counterproposes y, where $y_2 \in (x_2+c_L, x_2+c_H)$. If this event occurs off the equilibrium path, then the notion of sequential equilibrium does not impose any restriction on Player 1's beliefs about whom she faces. However, we argue that it is unreasonable, after this event occurs, for Player 1 to believe that she faces Player 2_H. The reason is as follows. Had Player 2 accepted the proposal he would have obtained x_2. If Player 1 accepts his counterproposal y, then Player 2 receives y_2 with one period of delay, which, if he is 2_H, is worse for him than receiving x_2 immediately (since $y_2 < x_2 + c_H$). On the other hand, Player 2_L is better off receiving y_2 with one period of delay than x_2 immediately (since $y_2 > x_2 + c_L$).

This argument is compatible with the logic of some of the recent refinements of the notion of sequential equilibrium—in particular that of Grossman and Perry (1986). In the language suggested by Cho and Kreps (1987), Player 2_L, when rejecting x and proposing y, can make the following speech. "I am Player 2_L. If you believe me and respond optimally, then you will accept the proposal y. In this case, it is not worthwhile for Player 2_H to pretend that he is I since he prefers the agreement x in the previous period to the agreement y this period. On the other hand it *is* worthwhile for me to persuade you that I am Player 2_L since I prefer the agreement y this period to the agreement x in the previous period. Thus, you should believe that I am Player 2_L."

Now suppose that Player 2 rejects the proposal x and counterproposes y, where $y_2 > x_2 + c_H$. In this case *both* types of Player 2 are better off if the counterproposal is accepted than they would have been had they accepted x, so that Player 1 has no reason to change the probability that she assigns to the event that she faces Player 2_H.

Thus we restrict attention to beliefs that are of the following form.

Definition 5.5 The beliefs of Player 1 are *rationalizing* if, after any history h for which $p_H(h) < 1$, they satisfy the following conditions.

1. If Player 2 rejects the proposal x and counteroffers y where $y_2 \in (x_2 + c_L, x_2 + c_H)$, then Player 1 assigns probability one to the event that she faces Player 2_L.
2. If Player 2 rejects the proposal x and counteroffers y where $y_2 > x_2 + c_H$, then Player 1's belief remains the same as it was before she proposed x.

We refer to a sequential equilibrium in which Player 1's beliefs are rationalizing as a *rationalizing sequential equilibrium.* The sequential equilibrium constructed in the proof of Part 3 of Proposition 5.3 is not rationalizing. If, for example, in state x^* of this equilibrium, Player 2 rejects a proposal x for which $x_1 > x_1^*$ and proposes y with $x_1 - c_H < y_1 < x_1 - c_L$, then the state changes to H, in which Player 1 believes that she faces Player 2_H with probability one. If Player 1 has rationalizing beliefs, however, she must believe that she faces Player 2_L with probability one in this case.

Lemma 5.6 *Every rationalizing sequential equilibrium of* $\Gamma(\pi_H)$ *has the following properties.*

1. *If Player* 2_H *accepts a proposal* x *for which* $x_1 > c_L$ *then Player* 2_L *rejects it and counterproposes* y*, with* $y_1 = \max\{0, x_1 - c_H\}$.
2. *Along the equilibrium path, agreement is reached in one of the following three ways.*
   a. *Players* 2_H *and* 2_L *make the same offer, which Player 1 accepts.*
   b. *Player 1 proposes* $(c_L, 1 - c_L)$*, which Players* 2_H *and* 2_L *both accept.*
   c. *Player 1 proposes* x *with* $x_1 \geq c_L$*, Player* 2_H *accepts this offer, and Player* 2_L *rejects it and proposes* y *with* $y_1 = \max\{0, x_1 - c_H\}$.
3. *If Player 1's payoff exceeds* $M_1 - 2c_1$*, where* M_1 *is the supremum of her payoffs over all rationalizing sequential equilibria, then agreement is reached immediately with Player* 2_H.

Proof. We establish each part separately.

1. Suppose that Player 2_H accepts the proposal x, for which $x_1 > c_L$. By Lemma 5.2, Player 2_L's strategy calls for him either to accept x or to

reject it and to counterpropose y with $\max\{0, x_1 - c_H\} \leq y_1 \leq x_1 - c_L$. In any case in which his strategy does not call for him to reject x and to propose y with $y_1 = \max\{0, x_1 - c_H\}$ he can deviate profitably by rejecting x and proposing z satisfying $\max\{0, x_1 - c_H\} < z_1 < y_1$. Upon seeing this counteroffer Player 1 accepts z since she concludes that she is facing Player 2_L.

2. Since in equilibrium Player 1 never proposes an agreement in which she gets less than c_L, the result follows from Lemma 5.2 and Part 1.

3. Consider an equilibrium in which Player 2_H rejects Player 1's initial proposal of x. By Lemma 5.2, Player 2_L also rejects this offer, and he and Player 2_H make the same counterproposal, say y. If Player 1 rejects y then her payoff is at most $M_1 - 2c_1$. If she accepts it, then her payoff is $y_1 - c_1$. Since Player 2_H rejected x in favor of y we must have $y_2 \geq x_2 + c_H$. Now, in order to make unprofitable the deviation by either of the types of Player 2 of proposing z with $z_2 > y_2$, Player 1 must reject such a proposal. If she does so, then by the condition that her beliefs be rationalizing and the fact that $y_2 \geq x_2 + c_H$, her belief does not change, so that play proceeds into $\Gamma(\pi_H)$. In order to make her rejection optimal, there must therefore be a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ in which her payoff is at least $y_1 + c_1$. Thus in any rationalizing sequential equilibrium in which agreement with Player 2_H is not reached immediately, Player 1's payoff is at most $M_1 - 2c_1$. □

We now establish the main result of this section.

Proposition 5.7 *For all $0 < \pi_H < 1$ the game $\Gamma(\pi_H)$ has a rationalizing sequential equilibrium, and every such equilibrium satisfies the following.*

1. *If $\pi_H > 2c_1/(c_1 + c_H)$ then the outcome is agreement in period 0 on $(1, 0)$ if Player 2 is 2_H, and agreement in period 1 on $(1 - c_H, c_H)$ if Player 2 is 2_L.*
2. *If $(c_1 + c_L)/(c_1 + c_H) < \pi_H < 2c_1/(c_1 + c_H)$ then the outcome is agreement in period 0 on $(c_H, 1 - c_H)$ if Player 2 is 2_H, and agreement in period 1 on $(0, 1)$ if Player 2 is 2_L.*
3. *If $\pi_H < (c_1 + c_L)/(c_1 + c_H)$ then the outcome is agreement in period 0 on $(c_L, 1 - c_L)$, whatever Player 2's type is.*

Proof. Let M_1 be the supremum of Player 1's payoffs in all rationalizing sequential equilibria of $\Gamma(\pi_H)$.

Step 1. If $\pi_H > 2c_1/(c_1 + c_H)$ then $\Gamma(\pi_H)$ has a rationalizing sequential equilibrium, and the outcome in every such equilibrium is that specified in Part 1 of the proposition.

		$*$	L
1	proposes	$(1,0)$	$(c_L, 1-c_L)$
	accepts	$x_1 \geq 1-c_H$	$x_1 \geq 0$
	belief	π_H	0
2_H	proposes	$(\max\{0, x_1-c_H\}, \min\{1, x_2+c_H\})$, where x is the offer just rejected	$(0,1)$
	accepts	$x_1 \leq 1$	$x_1 \leq c_H$
2_L	proposes	$(\max\{0, x_1-c_H\}, \min\{1, x_2+c_H\})$, where x is the offer just rejected	$(0,1)$
	accepts	$x_1 \leq c_L$	$x_1 \leq c_L$
Transitions		Go to L if Player 2 rejects x and counterproposes y with $y_2 \leq x_2 + c_H$.	Absorbing

Table 5.3 A rationalizing sequential equilibrium of $\Gamma(\pi_H)$ when $\pi_H \geq 2c_1/(c_1+c_H)$.

Proof. It is straightforward to check that the equilibrium described in Table 5.3 is a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ when $\pi_H > 2c_1/(c_1+c_H)$. (Note that state L is the same as it is in the sequential equilibria constructed in Section 5.3.) To establish the remainder of the claim, note that by Parts 2 and 3 of Lemma 5.6 we have $M_1 \leq \max\{\pi_H + (1-\pi_H)(1-c_H-c_1), c_L\}$. Under our assumption that $c_1 + c_L + c_H \leq 1$ we thus have $M_1 \leq \pi_H + (1-\pi_H)(1-c_H-c_1)$, and hence, by Part 1 of Proposition 5.3, Player 1's payoff in all rationalizing sequential equilibria is $\pi_H + (1-\pi_H)(1-c_H-c_1)$. The result follows from Part 2 of Lemma 5.6.

Step 2. If $\pi_H < 2c_1/(c_1+c_H)$ then $M_1 \leq \max\{\pi_H c_H + (1-\pi_H)(-c_1), c_L\}$.

Proof. Assume to the contrary that $M_1 > \max\{\pi_H c_H + (1-\pi_H)(-c_1), c_L\}$, and consider a rationalizing sequential equilibrium in which Player 1's payoff exceeds $M_1 - \epsilon > \max\{\pi_H c_H + (1-\pi_H)(-c_1), c_L\}$ for $0 < \epsilon < 2c_1 - \pi_H(c_1+c_H)$. By Part 3 of Lemma 5.6, Player 2_H accepts Player 1's offer x in period 0 in this equilibrium. By Parts 2b and 2c of the lemma it follows that $x_1 > c_H$ (and Player 2_L rejects x). We now argue that if Player 2_H deviates by rejecting x and proposing $z = (x_1 - c_H - \eta, x_2 + c_H + \eta)$ for some sufficiently small $\eta > 0$, then Player 1 accepts z, so that the deviation is profitable. If Player 1 rejects z, then, since her beliefs are unchanged (by the second condition in Definition 5.5), the most she can get is M_1 with a period of delay. But $x_1 - c_H - \epsilon > \pi_H(x_1 - c_1) + (1-\pi_H)(x_1 - c_H - 2c_1) \geq$

		$*$	L
1	proposes	z^*	$(c_L, 1 - c_L)$
	accepts	$x_1 \geq 0$	$x_1 \geq 0$
	belief	π_H	0
2_H	proposes	$(0, 1)$	$(0, 1)$
	accepts	$x_1 \leq c_H$	$x_1 \leq c_H$
2_L	proposes	$(0, 1)$	$(0, 1)$
	accepts	$x_1 \leq c_L$	$x_1 \leq c_L$
Transitions		Go to L if Player 2 rejects x and counterproposes y with $y_2 \leq x_2 + c_H$.	Absorbing

Table 5.4 A rationalizing sequential equilibrium of $\Gamma(\pi_H)$. When $z^* = (c_H, 1 - c_H)$ this is a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ for $(c_1 + c_L)/(c_1 + c_H) \leq \pi_H \leq 2c_1/(c_1 + c_H)$, and when $z^* = (c_L, 1 - c_L)$ it is a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ for $\pi_H \leq (c_1 + c_L)/(c_1 + c_H)$.

$M_1 - c_1 - \epsilon$ (the first inequality by the condition on ϵ, the second by the fact that Player 1's payoff in the equilibrium exceeds $M_1 - \epsilon$), so that for η small enough we have $x_1 - c_H - \eta > M_1 - c_1$. Hence Player 1 must accept z, making Player 2_H's deviation profitable. Thus there is no rationalizing sequential equilibrium in which Player 1's payoff exceeds $\pi_H c_H + (1 - \pi_H)(-c_1)$.

Step 3. If $(c_1 + c_L)/(c_1 + c_H) \leq \pi_H \leq 2c_1/(c_1 + c_H)$ then $\Gamma(\pi_H)$ has a rationalizing sequential equilibrium, and $M_1 \geq \pi_H c_H + (1 - \pi_H)(-c_1)$.

Proof. This follows from the fact that, for $z^* = (c_H, 1 - c_H)$ and $(c_1 + c_L)/(c_1 + c_H) \leq \pi_H \leq 2c_1/(c_1 + c_H)$, the equilibrium given in Table 5.4 is a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ in which Player 1's payoff is precisely $\pi_H c_H + (1 - \pi_H)(-c_1)$. (Note that when $z^* = (c_H, 1 - c_H)$ the players' actions are the same in state $*$ as they are in state x^* of the equilibrium in Part 3 of Proposition 5.3, for $x^* = (c_H, 1 - c_H)$; also, state L is the same as in that equilibrium.)

Step 4. If $(c_1 + c_L)/(c_1 + c_H) < \pi_H < 2c_1/(c_1 + c_H)$, then the outcome in every rationalizing sequential equilibrium is that specified in Part 2 of the proposition.

Proof. From Steps 2 and 3 we have $M_1 = \pi_H c_H + (1 - \pi_H)(-c_1)$. Since Player 2_H accepts any proposal in which Player 1 receives less than c_H, it follows that Player 1's expected payoff in all rationalizing sequential equilibria is precisely $\pi_H c_H + (1 - \pi_H)(-c_1)$. Given Lemma 5.6 this payoff can be obtained only if Player 1 proposes $(c_H, 1 - c_H)$, which Player 2_H accepts and Player 2_L rejects, and Player 2_L counterproposes $(0, 1)$, which Player 1 accepts.

Step 5. If $\pi_H \leq (c_1 + c_L)/(c_1 + c_H)$ then $\Gamma(\pi_H)$ has a rationalizing sequential equilibrium, and $M_1 \geq c_L$.

Proof. This follows from the fact that, for $z^* = (c_L, 1 - c_L)$ and $\pi_H \leq (c_1 + c_L)/(c_1 + c_H)$, the equilibrium given in Table 5.4 is a rationalizing sequential equilibrium of $\Gamma(\pi_H)$ in which Player 1's payoff is c_L.

Step 6. If $\pi_H < (c_1 + c_L)/(c_1 + c_H)$ then the outcome in every rationalizing sequential equilibrium is that specified in Part 3 of the proposition.

Proof. From Steps 2 and 5 we have $M_1 = c_L$. Since both types of Player 2 accept any proposal in which Player 1 receives less than c_L, it follows that Player 1's expected payoff in all rationalizing sequential equilibria is precisely c_L. The result follows from Part 2 of Lemma 5.6. □

The restriction on beliefs that is embedded in the definition of a rationalizing sequential equilibrium has achieved the target of isolating a unique outcome. However, the rationale for the restriction is dubious. First, the logic of the refinement assumes that Player 1 tries to rationalize any deviation of Player 2. If Player 2 rejects the offer x and makes a counteroffer in which his share is between $x_2 + c_L$ and $x_2 + c_H$, then Player 1 is assumed to interpret it as a signal that he is Player 2_L. However, *given the equilibrium strategies*, Player 2 does not benefit from such a deviation, so that another valid interpretation is that Player 2 is simply irrational. Second, if indeed Player 2 believes that it is possible to persuade Player 1 that he is Player 2_L by deviating in this way, then it seems that he should be regarded as irrational if he does not make the deviation that gives him the highest possible payoff (i.e. that in which his share is $x_2 + c_H$). Nevertheless, our refinement assumes that Player 1 interprets *any* deviation in which Player 2 counteroffers z with $z_2 \in (x_2 + c_L, x_2 + c_H)$ as a signal that Player 2 is Player 2_L. Thus we should be cautious in evaluating the result (and any other result that depends on a similar refinement of sequential equilibrium). In the literature on refinements of sequential equilibrium (see for example Cho and Kreps (1987) and van Damme (1987)) numerous restrictions on the beliefs are suggested, but none appears to generate a persuasive general criterion for selecting equilibria.

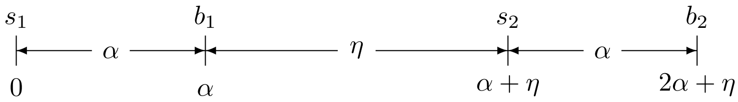


Figure 5.2 The reservation values of buyers and sellers.

5.6 Mechanism Design

In this section we depart from the study of sequential models and introduce some of the central ideas from the enormous literature on "mechanism design". We discuss only some ideas that are relevant to the analysis of bargaining between incompletely informed players; we do not provide a comprehensive introduction to the literature.

The study of mechanism design has two aims. The first is to design mechanisms that have desirable properties as devices for implementing outcomes in social conflicts. A discussion of the theory from this angle is beyond the scope of this book. The second aim is related to the criticism that strategic models of bargaining are too specific, since they impose a rigid structure on the bargaining process. The work on mechanism design provides a framework within which it is possible to analyze *simultaneously* a large set of bargaining procedures. A theory of bargaining is viewed as a mechanism that assigns an outcome to every possible configuration of the parameters of the model. A study of the set of mechanisms that can be generated by the Nash equilibria of bargaining games between incompletely informed players sheds light on the properties shared by these equilibria.

We focus on the following bargaining problem. A seller and a buyer of an indivisible good are negotiating a price. If they fail to reach agreement, each can realize a certain "reservation value". The reservation value s of the seller takes one of the two possible values s_1 and s_2, each with probability $1/2$; we refer to a seller with reservation value s_i as S_i. Similarly, the reservation value b of the buyer takes one of the two possible values b_1 and b_2, each with probability $1/2$; we refer to a buyer with reservation value b_j as B_j. The realizations of s and b are independent, so that all four combinations of s_i and b_j are equally likely. We assume that $s_1 < b_1 < s_2 < b_2$. To simplify the calculations we further restrict attention to the symmetric case in which $b_2 - s_2 = b_1 - s_1 = \alpha$; we let $s_2 - b_1 = \eta$, and (without loss of generality) let $s_1 = 0$. The reservation values are shown in Figure 5.2. Notice that the model departs from those of the previous sections in assuming that *both* bargainers are incompletely informed.

The tension in this bargaining problem is twofold. First, there is the usual conflict between a seller who is interested in obtaining a high price and a buyer who would like to pay as little as possible. Second, there is an incentive for B_2 to pretend to be B_1, thereby strengthening his bargaining position; similarly, there is an incentive for S_1 to pretend to be S_2.

A *mechanism* is a function that assigns an outcome to every realization of (s, b). To complete the definition we need to specify the set of possible outcomes. We confine attention to the case in which an *outcome* is a pair consisting of a price and a time at which the good is exchanged. Thus formally a mechanism is a pair (p, θ) of functions; p assigns a price, and θ a time in $[0, \infty]$, to each realization of (s, b). The interpretation is that if the realization of (s, b) is (s_i, b_j), then agreement is reached on the price $p(s_i, b_j)$ at time $\theta(s_i, b_j)$. The case $\theta(s_i, b_j) = \infty$ corresponds to that in which no trade ever occurs. (In most of the literature on mechanism design an outcome is a pair (p, π) with the interpretation that the good is exchanged for the price p with probability π and there is disagreement with probability $1 - \pi$. The results of this section can easily be translated into this case. We have chosen an alternative framework in order to make a clear comparison with the sequential bargaining models which are the focus of this book.)

The agents maximize expected utility. The utility of S_i for the price $\overline{p}$ at $\overline{\theta}$ is $\delta^{\overline{\theta}}(\overline{p} - s_i)$, and the utility of B_j for the price $\overline{p}$ at $\overline{\theta}$ is $\delta^{\overline{\theta}}(b_j - \overline{p})$; if an agent does not trade his utility is zero.

Let $M = (p, \theta)$ be a mechanism, and let

$$U_M(s_i) = \mathrm{E}_b[\delta^{\theta(s_i,b)}(p(s_i, b) - s_i)]$$

for $i = 1$, 2, where E_b denotes the expectation with respect to b (which is a random variable). Thus $U_M(s_i)$ is the expected utility of S_i from participating in the mechanism M. Similarly let

$$U_M(b_j) = \mathrm{E}_s[\delta^{\theta(s,b_j)}(b_j - p(s, b_j))]$$

for $j = 1$, 2, the expected utility of B_j from participating in the mechanism M. We consider mechanisms that satisfy the following conditions.

IR (*Individual Rationality*) $U_M(s_i) \geq 0$ and $U_M(b_j) \geq 0$ for i, $j = 1$, 2.

Behind IR is the assumption that each agent has the option of not taking part in the mechanism.

IC (*Incentive Compatibility*) For $(i, h) = (1, 2)$ and $(i, h) = (2, 1)$ we have

$$U_M(s_i) \geq \mathrm{E}_b[\delta^{\theta(s_h,b)}(p(s_h, b) - s_i)],$$

and for $(j,k)=(1,2)$ and $(j,k)=(2,1)$ we have

$$U_M(b_j) \geq \mathrm{E}_s[\delta^{\theta(s,b_k)}(b_j - p(s,b_k))].$$

Behind IC is the assumption that each agent has the option of imitating an agent with a different reservation value.

The connection between behavior in a strategic model of bargaining and these two conditions is the following. Consider a bargaining game in extensive form in which every terminal node corresponds to an agreement on a certain price at a certain time. Assume that the game is independent of the realization of the types: the strategy sets of the different types of buyer, and of seller, are the same, and the outcome of bargaining is a function only of the strategies used by the seller and the buyer. Any function that selects a Nash equilibrium for each realization of (s,b) is a mechanism. The fact that a strategy pair is a Nash equilibrium means that neither player can increase his payoff by adopting a different strategy. In particular, neither player can increase his payoff by adopting the strategy used by a player with a different reservation value. Thus the mechanism induced by a selection of Nash equilibria satisfies IC. If in the bargaining game each player has the option of not transacting, then the induced mechanism also satisfies IR.

Let $\sigma(s_i) = \mathrm{E}_b[\delta^{\theta(s_i,b)}]$, and similarly let $\beta(b_j) = \mathrm{E}_s[\delta^{\theta(s,b_j)}]$. Then we can write two of the incentive compatibility constraints as

$$U_M(s_1) \;\geq\; U_M(s_2) + (s_2 - s_1)\sigma(s_2) \tag{5.1}$$

$$U_M(b_2) \;\geq\; U_M(b_1) + (b_2 - b_1)\beta(b_1). \tag{5.2}$$

Our first observation concerns the existence of a mechanism (p,θ) that results in immediate agreement if the reservation value of the buyer exceeds that of the seller, and no transaction otherwise—i.e. in which $\theta(s_1,b_1) = \theta(s_1,b_2) = \theta(s_2,b_2) = 0$ and $\theta(s_2,b_1) = \infty$. We say that such a mechanism is *efficient*.

Proposition 5.8 *An efficient mechanism satisfying IR and IC exists if and only if* $s_2 - b_1 \leq (b_2 - s_2) + (b_1 - s_1)$ *(i.e. if and only if* $\eta \leq 2\alpha$*).*

Proof. We first show that if $\eta > 2\alpha$ then no efficient mechanism exists. The idea of the proof is that if η is large, then there is not enough surplus available to give S_1 a payoff high enough that she cannot benefit from imitating S_2. For any efficient mechanism $M = (p,\theta)$ we have $\sigma(s_2) = \beta(b_1) = 1/2$. If M satisfies IR and IC, then from (5.1) and (5.2) we have $U_M(s_1) \geq (s_2 - s_1)/2$ and $U_M(b_2) \geq (b_2 - b_1)/2$. Now, since the seller has reservation value s_1 with probability $1/2$, and the buyer has reservation value b_2 with probability $1/2$, the sum of the expected utilities of the seller

	b_1	b_2
s_1	b_1	$(b_2+s_1)/2$
s_2	–	s_2

Table 5.5 The price function for an efficient mechanism. The entry in the box (s_i, b_j) is the price $p(s_i, b_j)$. Since $\theta(s_2, b_1) = \infty$ in an efficient mechanism, $p(s_2, b_1)$ is irrelevant.

and the buyer is at least $U_M(s_1)/2 + U_M(b_2)/2$. From the above argument, this is equal to at least $(s_2 - s_1 + b_2 - b_1)/4 = (\alpha+\eta)/2$. But no transaction can generate a sum of utilities in excess of $[(b_2 - s_1) + (b_2 - s_2) + (b_1 - s_1) + 0]/4 = (4\alpha + \eta)/4$, which is strictly less than $(\alpha + \eta)/2$ if $\eta > 2\alpha$. Thus no efficient mechanism exists.

We now exhibit an efficient mechanism in the case that $\eta \leq 2\alpha$. The prices $p(s, b)$ specified by this mechanism are given in Table 5.5. It is straightforward to check that $p(s_i, b_j) \geq s_i$ and $b_j \geq p(b_j, s_i)$ for all i, j for which agreement is reached, so that the mechanism satisfies IR. To see that it satisfies IC, note that $U_M(s_1) = (2b_1 + b_2 - 3s_1)/4 = \alpha + \eta/4$, while S_1's utility if she pretends to be S_2 is $(s_2 - s_1)/2 = (\alpha + \eta)/2$. Since $\eta \leq 2\alpha$, the latter cannot exceed the former. The prices $p(s_1, b_j)$ for $j = 1$, 2, are both less than s_2, so that S_2 cannot benefit by imitating S_1. Symmetric arguments show that neither B_1 nor B_2 can benefit from imitating each other. □

Thus if $\eta > 2\alpha$ then in every mechanism some of the gains from trade are not exploited. What is the maximal sum of utilities in this case? We give an answer to this question for a restricted class of mechanisms.

Consider a bargaining game in which each player can unilaterally enforce disagreement (that is, he can refuse to participate in a trade from which he loses), the bargaining powers of the players are equal, and the bargaining procedure treats sellers and buyers symmetrically. A mechanism defined by a selection of symmetric Nash equilibria of such a game satisfies the following conditions.

IR* (*Ex Post Individual Rationality*) $s_i \leq p(s_i, b_j) \leq b_j$ whenever $\theta(s_i, b_j) < \infty$, and $\theta(s_2, b_1) = \infty$.

This condition says that no agreement is reached if the buyer's reservation value is smaller than the seller's, and that both parties to an agreement must benefit *after* their identities are determined. Note that IR involves

a player's decision to participate in the mechanism *before* he is aware of the realization of his opponent's reservation value, while IR* involves his decision to trade after this realization. Obviously, IR* implies IR.

SY (*Symmetry*) $p(s_1, b_2) = (s_1 + b_2)/2 = \alpha + \eta/2$, $b_2 - p(s_2, b_2) = p(s_1, b_1)$, and $\theta(s_2, b_2) = \theta(s_1, b_1)$.

This condition expresses the symmetry between a buyer with a high reservation value and a seller with a low reservation value, as well as that between a seller with a high reservation value and a buyer with a low reservation value. It requires that in the bargaining between S_1 and B_2 the surplus be split equally, that the time of trade between B_2 and S_2 is the same as that between S_1 and B_1, and that the utilities obtained by S_1 and B_2 are the same.

The conditions IR* and SY reduce the choice of a mechanism to the choice of a triple $(p(s_1, b_1), \theta(s_1, b_1), \theta(s_1, b_2))$. (Note that $p(s_2, b_1)$ is irrelevant since $\theta(s_2, b_1) = \infty$.) Since $p(s_1, b_2) < s_2$, S_2 cannot gain by imitating S_1, and similarly B_1 cannot gain by imitating B_2. The condition that S_1 not benefit from imitating S_2 is $\delta^{\theta(s_1,b_2)}[(s_1+b_2)/2] + \delta^{\theta(s_1,b_1)} p(s_1, b_1) \geq \delta^{\theta(s_2,b_2)} p(s_2, b_2) = \delta^{\theta(s_1,b_1)}(b_2 - p(s_1, b_1))$, which is equivalent to $\delta^{\theta(s_1,b_2)}(\alpha + \eta/2) \geq \delta^{\theta(s_1,b_1)}(2\alpha + \eta - 2p(s_1, b_1))$. The sum of the expected utilities is $\delta^{\theta(s_1,b_1)}\alpha/2 + \delta^{\theta(s_1,b_2)}(2\alpha + \eta)/4$. This is maximized, subject to the constraint, by $\theta(s_1, b_2) = 0$, $p(s_1, b_1) = \alpha$, and $\delta^{\theta(s_1,b_1)} = (\alpha + \eta/2)/\eta = 1/2 + \alpha/\eta$. (Note that $\delta^{\theta(s_1,b_1)} < 1$ since $\eta > 2\alpha$.) We have proved the following.

Proposition 5.9 *If $\eta > 2\alpha$ then the mechanism, among those that satisfy IC, SY, and IR*, that maximizes the sum of the expected utilities is given by the following: $\delta^{\theta(s_1,b_1)} = \delta^{\theta(s_2,b_2)} = 1/2 + \alpha/\eta$, $\theta(b_1, s_2) = \infty$, $\theta(s_1, b_2) = 0$, $p(s_1, b_1) = \alpha$, $p(s_2, b_2) = \alpha + \eta$, and $p(s_1, b_2) = (s_1 + b_2)/2$. The minimized loss of expected utilities is $(1/2 - \alpha/\eta)\alpha/2 = \alpha/4 - \alpha^2/2\eta$.*

This result gives us a lower bound on the inefficiency that is unavoidable in the outcome of any bargaining game with incomplete information. The following is a game in which the Nash equilibrium outcome induces the mechanism described in the proposition, thus showing that the lower bound can be achieved. Each of the players has to announce a type: the strategy set of the seller is $\{s_1, s_2\}$, and that of the buyer is $\{b_1, b_2\}$. The outcomes of the four possible strategy choices are given in Table 5.6, where $\delta^{\theta^*} = 1/2 + \alpha/\eta$. This game has a Nash equilibrium in which each player announces his true type; the outcome is the one described in Proposition 5.9. However, the game is highly artificial; we fail to see any "natural" game that induces the mechanism described in the proposition. In particular,

	b_1	b_2
s_1	price α at time θ^*	price $(b_2 + s_1)/2$ at time 0
s_2	disagreement	price $\alpha + \eta$ at time θ^*

Table 5.6 The outcomes in a game for which the Nash equilibrium minimizes the loss of surplus. The time θ^* is defined by $\delta^{\theta^*} = 1/2 + \alpha/\eta$.

in any game that does so the players must be prevented from renegotiating the date of agreement in those cases in which delayed agreement is prescribed.

To summarize, in this section we have described some representative results from the literature on mechanism design. Proposition 5.8 shows that inefficiency is inherent in the outcomes of a large family of bargaining games with incomplete information, while Proposition 5.9 characterizes the minimal loss of the sum of the utilities that is associated with these games. Note that these results apply to games in which the players have a particular type of preferences. They are not immediately applicable to other cases, like those in which the players have different discount factors or fixed bargaining costs. In these cases, the set of mechanisms that satisfy IR and IC may be larger than the results in this section suggest. Alternatively, we may wish to restrict the class of mechanisms that we consider: for example, we may wish to take into account the possibility that the players will renegotiate if the mechanism assigns a delayed agreement. Note further that the fact that an efficient mechanism exists does not mean that it is plausible. For example, in the case in which only one player—say the seller—is uncertain of her opponent's type, and $s < b_1 < b_2$, the mechanism design problem is trivial. For every price $\overline{p}$ between s and b_1, the mechanism (p, θ) in which $p(s, b_i) = \overline{p}$ and $\theta(s, b_i) = 0$ for $i = 1$, 2 is an efficient mechanism that satisfies IC and IR. Nevertheless, the outcome of reasonable bargaining games (like those described in earlier sections) may be inefficient: the fact that the incentive compatibility constraints alone do not imply that there must be inefficiency does not mean that the outcome of an actual game will be efficient.

Notes

Sections 5.2 and 5.3 (which discuss the basic alternating offers model with one-sided uncertainty) are based on Rubinstein (1985a, b). The discus-

sion in Section 5.4 of the delay in reach an agreement when the strategies of Players 2_H and 2_L are stationary originated in Gul and Sonnenschein (1988). Section 5.5, in which we study a refinement of sequential equilibrium, is based on Rubinstein (1985a, b). The discussion of mechanism design in Section 5.6 originated in Myerson and Satterthwaite (1983); our treatment is based on Matsuo (1989).

We have not considered in this chapter the axiomatic approach to bargaining with incomplete information. A paper of particular note in this area is Harsanyi and Selten (1972), who extend the Nash bargaining solution to the case in which the players are incompletely informed.

The literature on bargaining between incompletely informed players is very large. We mention only a small sample here. A collection of papers in the field is Roth (1985).

In the strategic models we have studied in this chapter, only *one* of the players is incompletely informed. Cramton (1992) constructs a sequential equilibrium for a bargaining game of alternating offers in which both players are incompletely informed. Ausubel and Deneckere (1992a), Chatterjee and Samuelson (1988), and Cho (1989) further analyze this case.

Admati and Perry (1987) study a model in which a player who rejects an offer chooses how long to wait before making a counteroffer. By halting the negotiations for some time, a player can persuade his opponent that he is a "strong" bargainer, to whom concessions should be made. Thus the delay is a signal that reveals information. In this model, in some cases, the delay before an agreement is reached remains significant even when the length of each period converges to zero.

Among other papers that analyze strategic models of bargaining in which the players alternate offers are the following. Bikhchandani (1986) investigates the assumption that Player 1 updates her beliefs after Player 2 responds, *before* he makes a counteroffer. Bikhchandani (1992) explores the consequences of imposing different restrictions on players' beliefs in events that do not occur if the players follow their equilibrium strategies. Grossman and Perry (1986) propose a refinement of sequential equilibrium in the case in which there are many possible types (not just two) for Player 2. Perry (1986) studies a model in which the proposer is determined endogenously in the game. Sengupta and Sengupta (1988) consider a model in which an offer is a contract that specifies a division of the pie contingent on the state. Ausubel and Deneckere (1992b) further study the model of Gul and Sonnenschein (1988) (see the end of Section 5.4). In a model like that of Gul and Sonnenschein (1988), Vincent (1989) demonstrates that if the seller's and buyer's values for the good are correlated, then delay is possible in a sequential equilibrium that satisfies conditions like those of Section 5.5 (see also Cothren and Loewenstein (n.d.)).

If we assume that there are only two possible agreements, then the complexity of the analysis is reduced dramatically, allowing sharp results to be established (see, for example, Chatterjee and Samuelson (1987) and Ponsati-Obiols (1989, 1992)). This literature is closely connected with that on wars of attrition (see, for example, Osborne (1985)), since accepting the worst agreement in a bargaining game is analogous to conceding in a war of attrition.

Some of the results in models of bargaining with one-sided incomplete information in which both parties make offers can be obtained also in models in which only the uninformed party is allowed to make offers. An extensive survey of this literature is given in Fudenberg, Levine, and Tirole (1985).

PART 2

Models of Decentralized Trade

The models of bargaining in Part 1 concern isolated encounters between pairs of players; the outcome in the event of disagreement is exogenous. We now study decentralized markets, in which many pairs of agents simultaneously bargain over the gains from trade. The outcome in any match depends upon events outside that match and upon the agents' expectations about these events. In particular, an agent's evaluation of a termination of the negotiation, whether this termination is a result of an exogenous event or a deliberate action by one of the parties, depends upon the outcome of negotiation between these agents and alternative partners. Further, the existence and identities of these alternative partners are affected by the outcome of negotiation between other pairs of agents. In short, in the models we study, the outcome of negotiation between any pair of agents may be influenced by the outcome in other bargaining encounters. The solution of one bargaining situation is part of an *equilibrium* in the entire market.

The models we study assist our understanding of the working of markets. For each model, we consider the relation of the outcome with the "Competitive Equilibrium". Our models indicate the scope of the competitive model: when it is appropriate, and when it is not. In case it is not, we investigate how the outcome depends on the time structure of trade and the information possessed by the traders.

CHAPTER 6

A First Approach Using the Nash Solution

6.1 Introduction

There are many choices to be made when constructing a model of a market in which individuals meet and negotiate prices at which to trade. In particular, we need to specify the process by which individuals are matched, the information that the individuals possess at each point in time, and the bargaining procedure that is in use. We consider a number of possibilities in the subsequent chapters. In most cases (the exception is the model in Section 8.4), we study a market in which the individuals are of two types: (potential) sellers and (potential) buyers. Each transaction takes place between a seller and a buyer, who negotiate the terms of the transaction.

In this chapter we use the Nash bargaining solution (see Chapter 2) to model the outcome of negotiation. In the subsequent chapters we model the negotiation in more detail, using strategic models like the one in Chapter 3.

We distinguish two possibilities for the evolution of the number of traders present in the market.

1. The market is in a steady state. The number of buyers and the number of sellers in the market remain constant over time. The

opportunities for trade remain unchanged. The pool of potential buyers may always be larger than the pool of potential sellers, but the discrepancy does not change over time. An example of what we have in mind is the market for apartments in a city in which the rate at which individuals vacate their apartments is similar to the rate at which individuals begin searching for an apartment.

2. All the traders are present in the market initially. Entry to the market occurs only once. A trader who makes a transaction in some period subsequently leaves the market. As traders complete transactions and leave the market, the number of remaining traders dwindles. When all possible transactions have been completed, the market closes. A periodic market for a perishable good is an example of what we have in mind.

In Sections 6.3 and 6.4 we study models founded on these two assumptions. The primitives in each model are the numbers of traders present in the market. Alternatively we can construct models in which these numbers are determined endogenously. In Section 6.6 we discuss two models based on those in Sections 6.3 and 6.4 in which each trader decides whether or not to enter the market. The primitives in these models are the numbers of traders *considering* entering the market.

6.2 Two Basic Models

In this section we describe two models, in which the number of traders in the market evolves in the two ways discussed above. Before describing the differences between the models, we discuss features they have in common.

Goods A single indivisible good is traded for some quantity of a divisible good ("money").

Time Time is discrete and is indexed by the integers.

Economic Agents Two types of agents operate in the market: "sellers" and "buyers". Each seller owns one unit of the indivisible good, and each buyer owns one unit of money. Each agent concludes at most one transaction. The characteristics of a transaction that are relevant to an agent are the price p and the number of periods t after the agent's entry into the market that the transaction is concluded. Each individual's preferences on lotteries over the pairs (p, t) satisfy the assumptions of von Neumann and Morgenstern. Each seller's preferences are represented by the utility function $\delta^t p$, where $0 <$

$\delta \leq 1$, and each buyer's preferences are represented by the utility function $\delta^t(1-p)$. If an agent never trades then his utility is zero.

The roles of buyers and sellers are symmetric. The only asymmetry is that the numbers of sellers and buyers in the market at any time may be different.

Matching Let B and S be the numbers of buyers and sellers active in some period t. Every agent is matched with at most one agent of the opposite type. If $B > S$ then every seller is matched with a buyer, and the probability that a buyer is matched with some seller is equal to S/B. If sellers outnumber buyers, then every buyer is matched with a seller, and a seller is matched with a buyer with probability B/S. In both cases the probability that any given pair of traders are matched is independent of the traders' identities.

This matching technology is special, but we believe that most of the results below can be extended to many other matching technologies.

Bargaining When matched in some period t, a buyer and a seller negotiate a price. If they do not reach an agreement, each stays in the market until period $t+1$, when he has the chance of being matched anew. If there exists no agreement that both prefer to the outcome when they remain in the market till period $t+1$, then they do not reach an agreement. Otherwise in period t they reach the agreement given by the Nash solution of the bargaining problem in which a utility pair is feasible if it arises from an agreement concluded in period t, and the disagreement utility of each trader is his expected utility if he remains in the market till period $t+1$.

Note that the expected utility of an agent staying in the market until period $t+1$ may depend upon whether other pairs of agents reach agreement in period t.

We saw in Chapter 4 (see in particular Section 4.6) that the disagreement point should be chosen to reflect the forces that drive the bargaining process. By specifying the utility of an agent in the event of disagreement to be the value of being a trader in the next period, we are thinking of the Nash solution in terms of the model in Section 4.2. That is, the main pressure on the agents to reach an agreement is the possibility that negotiation will break down.

The differences between the models we analyze concern the evolution of the number of participants over time.

Model A The numbers of sellers and buyers in the market are constant over time.

A literal interpretation of this model is that a new pair consisting of a buyer and a seller springs into existence the moment a transaction is completed. Alternatively, we can regard the model as an approximation for the case in which the numbers of buyers and sellers are roughly constant, any fluctuations being small enough to be ignored by the agents.

Model B All buyers and sellers enter the market simultaneously; no new agents enter the market at any later date.

6.3 Analysis of Model A (A Market in Steady State)

Time runs indefinitely in both directions: the set of periods is the set of all integers, positive and negative. In every period there are S_0 sellers and B_0 buyers in the market. Notice that the primitives of the model are the *numbers* of buyers and sellers, not the sets of these agents. Sellers and buyers are not identified by their names or by their histories in the market. An agent is characterized simply by the fact that he is interested either in buying or in selling the good.

We restrict attention to situations in which all matches in all periods result in the same outcome. Thus, a candidate p for an equilibrium is either a price (a number in $[0,1]$), or D, the event that no agreement is reached. We denote the expected utilities of being a seller and a buyer in the market by V_s and V_b, respectively.

Given the linearity of the traders' utility functions in price, the set of utility pairs feasible within any given match is

$$U = \{(u_s, u_b) \in \mathcal{R}^2 : u_s + u_b = 1 \text{ and } u_i \geq 0 \text{ for } i = s, b\}. \tag{6.1}$$

If in period t a seller and buyer fail to reach an agreement, they remain in the market until period $t+1$, at which time their expected utilities are V_i for $i = s$, b. Thus from the point of view of period t, disagreement results in expected utilities of δV_i for $i = s$, b. So according to our bargaining solution, there is disagreement in any period if $\delta V_s + \delta V_b > 1$. Otherwise agreement is reached on the Nash solution of the bargaining problem $\langle U, d\rangle$, where $d = (\delta V_s, \delta V_b)$.

Definition 6.1 If $B_0 \geq S_0$ then an outcome p^* is a *market equilibrium* in Model A if there exist numbers $V_s \geq 0$ and $V_b \geq 0$ satisfying the following two conditions. First, if $\delta V_s + \delta V_b \leq 1$ then $p^* \in [0,1]$ and

$$p^* - \delta V_s = 1 - p^* - \delta V_b, \tag{6.2}$$

and if $\delta V_s + \delta V_b > 1$ then $p^* = D$. Second,

$$V_s = \begin{cases} p^* & \text{if } p^* \in [0,1] \\ \delta V_s & \text{if } p^* = D \end{cases} \tag{6.3}$$

and

$$V_b = \begin{cases} (S_0/B_0)(1-p^*) + (1-S_0/B_0)\delta V_b & \text{if } p^* \in [0,1] \\ \delta V_b & \text{if } p^* = D. \end{cases} \tag{6.4}$$

The first part of the definition requires that the agreement reached by the agents be given by the Nash solution. The second part defines the numbers V_i ($i = s$, b). If p^* is a price then $V_s = p^*$ (since a seller is matched with probability one), and $V_b = (S_0/B_0)(1-p^*) + (1-S_0/B_0)\delta V_b$ (since a buyer in period t is matched with probability S_0/B_0, and otherwise stays in the market until period $t+1$).

The definition for the case $B_0 \le S_0$ is symmetric. The following result gives the unique market equilibrium of Model A.

Proposition 6.2 *If $\delta < 1$ then there is a unique market equilibrium p^* in Model A. In this equilibrium agreement is reached and*

$$p^* = \begin{cases} \dfrac{1}{2-\delta+\delta S_0/B_0} & \textit{if } B_0 \ge S_0 \\ 1 - \dfrac{1}{2-\delta+\delta B_0/S_0} & \textit{if } B_0 \le S_0. \end{cases}$$

Proof. We deal only with the case $B_0 \ge S_0$ (the other case is symmetric). If $p^* = D$ then by (6.3) and (6.4) we have $V_s = V_b = 0$. But then agreement must be reached. The rest follows from substituting the values of V_s and V_b given by (6.3) and (6.4) into (6.2). □

The equilibrium price p^* has the following properties. An increase in S_0/B_0 decreases p^*. As the traders become more impatient (the discount factor δ decreases) p^* moves toward $1/2$. The limit of p^* as $\delta \to 1$ is $B_0/(S_0+B_0)$. (Note that if δ is *equal* to 1 then every price in $[0,1]$ is a market equilibrium.)

The primitives of the model are the numbers of buyers and sellers in the market. Alternatively, we can take the probabilities of buyers and sellers being matched as the primitives. If $B_0 > S_0$ then the probability of being matched is one for a seller and S_0/B_0 for a buyer. If we let these probabilities be the arbitrary numbers σ for a seller and β for a buyer (the same in every period), we need to modify the definition of a market equilibrium: (6.3) and (6.4) must be replaced by

$$V_s = \sigma p^* + (1-\sigma)\delta V_s \tag{6.5}$$

$$V_b = \beta(1-p^*) + (1-\beta)\delta V_b. \tag{6.6}$$

In this case the limit of the unique equilibrium price as $\delta \to 1$ is $\sigma/(\sigma+\beta)$.

The constraint that the equilibrium price not depend on time is not necessary. Extending the definition of a market equilibrium to allow the price on which the agents reach agreement to depend on t introduces no new equilibria.

6.4 Analysis of Model B (Simultaneous Entry of All Sellers and Buyers)

In Model B time starts in period 0, when S_0 sellers and B_0 buyers enter the market; the set of periods is the set of nonnegative integers. In each period buyers and sellers are matched and engage in negotiation. If a pair agrees on a price, the members of the pair conclude a transaction and leave the market. If no agreement is reached, then both individuals remain in the market until the next period. No more agents enter the market at any later date. As in Model A the primitives are the numbers of sellers and buyers in the market, not the sets of these agents.

A candidate for a market equilibrium is a function p that assigns to each pair (S, B) either a price in $[0, 1]$ or the disagreement outcome D. In any given period, the same numbers of sellers and buyers leave the market, so that we can restrict attention to pairs (S, B) for which $S \leq S_0$ and $B - S = B_0 - S_0$. Thus the equilibrium price may depend on the numbers of sellers and buyers in the market but not on the period. Our working assumption is that buyers initially outnumber sellers ($B_0 > S_0$).

Given a function p and the matching technology we can calculate the expected value of being a seller or a buyer in a market containing S sellers and B buyers. We denote these values by $V_s(S, B)$ and $V_b(S, B)$, respectively. The set of utility pairs feasible in any given match is U, as in Model A (see (6.1)). The number of traders in the market may vary over time, so the disagreement point in any match is determined by the equilibrium. If $p(S, B) = D$ then all the agents in the market in period t remain until period $t+1$, so that the utility pair in period $t+1$ is $(\delta V_s(S, B), \delta V_b(S, B))$. If at the pair (S, B) there is agreement in equilibrium (i.e. $p(S, B)$ is a price), then if any one pair fails to agree there will be one seller and $B - S + 1$ buyers in the market at time $t + 1$. Thus the disagreement point in this case is $(\delta V_s(1, B - S + 1), \delta V_b(1, B - S + 1))$. An appropriate definition of market equilibrium is thus the following.

Definition 6.3 If $B_0 \geq S_0$ then a function p^* that assigns an outcome to each pair (S, B) with $S \leq S_0$ and $S - B = S_0 - B_0$ is a *market equilibrium* in Model B if there exist functions V_s and V_b with $V_s(S, B) \geq 0$ and $V_b(S, B) \geq 0$ for all (S, B), satisfying the following two conditions. First, if $p^*(S, B) \in$

$[0,1]$ then $\delta V_s(1, B-S+1) + \delta V_b(1, B-S+1) \leq 1$ and

$$p^*(S,B) - \delta V_s(1, B-S+1) = 1 - p^*(S,B) - \delta V_b(1, B-S+1), \quad (6.7)$$

and if $p^*(S,B) = D$ then $\delta V_s(S,B) + \delta V_b(S,B) > 1$. Second,

$$V_s(S,B) = \begin{cases} p^*(S,B) & \text{if } p^*(S,B) \in [0,1] \\ \delta V_s(S,B) & \text{if } p^*(S,B) = D \end{cases} \quad (6.8)$$

and

$$V_b(S,B) = \begin{cases} (S/B)(1 - p^*(S,B)) & \text{if } p^*(S,B) \in [0,1] \\ \delta V_b(S,B) & \text{if } p^*(S,B) = D. \end{cases} \quad (6.9)$$

As in Definition 6.1, the first part ensures that the negotiated price is given by the Nash solution relative to the appropriate disagreement point. The second part defines the value of being a seller and a buyer in the market. Note the difference between (6.9) and (6.4). If agreement is reached in period t, then in the market of Model B no sellers remain in period $t+1$, so any buyer receives a payoff of zero in that period. Once again, the definition for the case $B_0 \leq S_0$ is symmetric. The following result gives the unique market equilibrium of Model B.

Proposition 6.4 *Unless $\delta = 1$ and $S_0 = B_0$, there is a unique market equilibrium p^* in Model B. In this equilibrium agreement is reached, and p^* is defined by*

$$p^*(S,B) = \begin{cases} \dfrac{1 - \delta/(B-S+1)}{2 - \delta - \delta/(B-S+1)} & \textit{if } B \geq S \\[2ex] \dfrac{1-\delta}{2 - \delta - \delta/(S-B+1)} & \textit{if } S \geq B. \end{cases}$$

Proof. We give the argument for the case $B_0 \geq S_0$; the case $B_0 \leq S_0$ is symmetric. We first show that $p^*(S,B) \neq D$ for all (S,B). If $p^*(S,B) = D$ then by (6.8) and (6.9) we have $V_i(S,B) = 0$ for $i = s, b$, so that $\delta V_s(S,B) + \delta V_b(S,B) \leq 1$, contradicting $p^*(S,B) = D$. It follows from (6.7) that the outcomes in markets with a single seller determine the prices upon which agreement is reached in all other markets. Setting $S = 1$ in (6.8) and (6.9), and substituting these into (6.7) we obtain

$$V_s(1,B) = \frac{2BV_s(1,B)}{\delta(B+1)} - \frac{B-\delta}{\delta(B+1)}.$$

This implies that $V_s(1,B) = (1 - \delta/B)/(2 - \delta - \delta/B)$. (The denominator is positive unless $\delta = 1$ and $B = 1$.) The result follows from (6.7), (6.8), and (6.9) for arbitrary values of S and B. □

The equilibrium price has properties different from those of Model A. In particular, if $S_0 < B_0$ then the limit of the price as $\delta \to 1$ (i.e. as the impatience of the agents diminishes) is 1. If $S_0 = B_0$ then $p^*(S, B) = 1/2$ for all values of $\delta < 1$. Thus the limit of the equilibrium price as $\delta \to 1$ is discontinuous as a function of the numbers of sellers and buyers.

As in Model A the constraint that the prices not depend on time is not necessary. If we extend the definition of a market equilibrium to allow p^* to depend on t in addition to S and B then no new equilibria are introduced.

6.5 A Limitation of Modeling Markets Using the Nash Solution

Models A and B illustrate an approach for analyzing markets in which prices are determined by bargaining. One of the attractions of this approach is its simplicity. We can achieve interesting insights into the agents' market interaction without specifying a strategic model of bargaining. However, the approach is not without drawbacks. In this section we demonstrate that it fails when applied to a simple variant of Model B.

Consider a market with one-time entry in which there is one seller whose reservation value is 0 and two buyers B_L and B_H whose reservation values are v_L and $v_H > v_L$, respectively. A candidate for a market equilibrium is a pair (p_H, p_L), where p_I is either a price (a number in $[0, v_H]$) or disagreement (D). The interpretation is that p_I is the outcome of a match between the seller and B_I. A pair (p_H, p_L) is a market equilibrium if there exist numbers V_s, V_L, and V_H that satisfy the following conditions. First

$$p_H = \begin{cases} \delta V_s + (v_H - \delta V_s - \delta V_H)/2 & \text{if } \delta V_s + \delta V_H \leq v_H \\ D & \text{otherwise} \end{cases}$$

and

$$p_L = \begin{cases} \delta V_s + (v_L - \delta V_s - \delta V_L)/2 & \text{if } \delta V_s + \delta V_L \leq v_L \\ D & \text{otherwise.} \end{cases}$$

Second, $V_s = V_H = V_L = 0$ if $p_H = p_L = D$; $V_s = (p_H + p_L)/2$, $V_I = (v_I - p_I)/2$ for $I = H, L$ if both p_H and p_L are prices; and $V_s = p_I/(2 - \delta)$, $V_I = (v_I - p_I)/(2 - \delta)$, and $V_J = 0$ if only p_I is a price.

If $v_H < 2$ and δ is close enough to one then this system has no solution.

In Section 9.2 we construct equilibria for a strategic version of this model. In these equilibria the outcome of a match is not independent of the history that precedes the match. Using the approach of this chapter we fail to find these equilibria since we implicitly restrict attention to cases in which the outcome of a match is independent of past events.

6.6 Market Entry

In the models we have studied so far, the primitive elements are the stocks of buyers and sellers present in the market. By contrast, in many markets agents can decide whether or not to participate in the trading process. For example, the owner of a good may decide to consume the good himself; a consumer may decide to purchase the good he desires in an alternative market. Indeed, economists who use the competitive model often take as primitive the characteristics of the traders who are considering entering the market.

6.6.1 Market Entry in Model A

Suppose that in each period there are $\overline{S}$ sellers and $\overline{B}$ buyers considering entering the market, where $\overline{B} > \overline{S}$. Those who do not enter disappear from the scene and obtain utility zero. The market operates as before: buyers and sellers are matched, conclude agreements determined by the Nash solution, and leave the market. We look for an equilibrium in which the numbers of sellers and buyers participating in the market are constant over time, as in Model A.

Each agent who enters the market bears a small cost $\epsilon > 0$. Let $V_i^*(S, B)$ be the expected utility of being an agent of type i $(= s,\ b)$ in a market equilibrium of Model A when there are $S > 0$ sellers and $B > 0$ buyers in the market; set $V_s^*(S, 0) = V_b^*(0, B) = 0$ for any values of S and B. If there are large numbers of agents of each type in the market, then the entry of an additional buyer or seller makes little difference to the equilibrium price (see Proposition 6.2). Assume that each agent believes that his own entry has no effect at all on the market outcome, so that his decision to enter a market containing S sellers and B buyers involves simply a comparison of ϵ with the value $V_i^*(S, B)$ of being in the market. (Under the alternative assumption that each agent anticipates the effect of his entry on the equilibrium, our main results are unchanged.)

It is easy to see that there is an equilibrium in which no agents enter the market. If there is no seller in the market then the value to a buyer of entering is zero, so that no buyer finds it worthwhile to pay the entry cost $\epsilon > 0$. Similarly, if there is no buyer in the market, then no seller finds it optimal to enter.

Now consider an equilibrium in which there are constant positive numbers S^* of sellers and B^* of buyers in the market at all times. In such an equilibrium a positive number of buyers (and an equal number of sellers) leaves the market each period. In order for these to be replaced by entering buyers we need $V_b^*(S^*, B^*) \geq \epsilon$. If $V_b^*(S^*, B^*) > \epsilon$ then all $\overline{B}$ buyers

contemplating entry find it worthwhile to enter, a number that needs to be balanced by sellers in order to maintain the steady state but cannot be even if all $\overline{S}$ sellers enter, since $\overline{B} > \overline{S}$. Thus in any steady state equilibrium we have $V_b^*(S^*, B^*) = \epsilon$.

If $S^* > B^*$ then by Proposition 6.2 we have $V_b^*(S^*, B^*) = 1/(2 - \delta + \delta B^*/S^*)$, so that $V_b^*(S^*, B^*) > 1/2$. Thus as long as $\epsilon < 1/2$ the fact that $V_b^*(S^*, B^*) = \epsilon$ implies that $S^* \leq B^*$. From Proposition 6.2 and (6.4) we conclude that

$$V_b^*(S^*, B^*) = \frac{S^*/B^*}{2 - \delta + \delta S^*/B^*},$$

so that $S^*/B^* = (2 - \delta)\epsilon/(1 - \delta\epsilon)$, and hence

$$p^* = V_s^*(S^*, B^*) = \frac{1 - \delta\epsilon}{2 - \delta}.$$

Thus $V_s^*(S^*, B^*) > \epsilon$, so that all $\overline{S}$ sellers enter the market each period. Active buyers outnumber sellers ($B^* > S^*$), so all S^* sellers leave the market every period. Hence $S^* = \overline{S}$, and $B^* = \overline{S}(1 - \delta\epsilon)/\epsilon(2 - \delta)$.

We have shown that in a nondegenerate steady state equilibrium in which the entry cost is small (less than 1/2) all $\overline{S}$ sellers enter the market each period, accompanied by the same number of buyers. All the sellers are matched, conclude an agreement, and leave the market. The constant number B^* of buyers in the market exceeds the number S^* of sellers. (For fixed δ, the limit of S^*/B^* as $\epsilon \to 0$ is zero.) The excess of buyers over sellers is just large enough to hold the value of being a buyer down to the (small) entry cost. Each period $\overline{S}$ of the buyers are matched, conclude an agreement, and leave the market. The remainder stay in the market until the next period, when they are joined by $\overline{S}$ new buyers.

The fact that $\delta < 1$ and $\epsilon > 0$ creates a "friction" in the market. As this friction converges to zero the equilibrium price converges to 1:

$$\lim_{\delta \to 1, \epsilon \to 0} p^* = 1.$$

In both Model A and the model of this section the primitives are numbers of sellers and buyers. In Model A, where these numbers are the numbers of sellers and buyers present in the market, we showed that if the number of sellers slightly exceeds the number of buyers then the limiting equilibrium price as $\delta \to 1$ is close to 1/2. When these numbers are the numbers of sellers and buyers considering entry into the market then this limiting price is 1 whenever the number of potential buyers exceeds the number of potential sellers.

6.6.2 Market Entry in Model B

Now consider the effect of adding an entry decision to Model B. As in the previous subsection, there are $\overline{S}$ sellers and $\overline{B}$ buyers considering entering the market, with $\overline{B} > \overline{S}$.

Each agent who enters bears a small cost $\epsilon > 0$. Let $V_i^*(S, B)$ be the expected utility of being an agent of type i $(= s, b)$ in a market equilibrium of Model B when $S > 0$ sellers and $B > 0$ buyers enter in period 0; set $V_s^*(S, 0) = V_b^*(0, B) = 0$ for any values of S and B.

Throughout the analysis we assume that the discount factor δ is close to 1. In this case the equilibrium price in Model B is very sensitive to the ratio of buyers to sellers: the entry of a single seller or buyer into a market in which the numbers of buyers and sellers are equal has a drastic effect on the equilibrium price (see Proposition 6.4). A consequence is that the agents' beliefs about the effect of their entry on the market outcome are critical in determining the nature of an equilibrium.

First maintain the assumption of the previous subsection that each agent takes the market outcome as given when deciding whether or not to enter. An agent of type i simply compares the expected utility $V_i^*(S, B)$ of an agent of his type currently in the market with the cost ϵ of entry. As before, there is an equilibrium in which no agent enters the market. However, in this case there are no other equilibria. To show this, first consider the possibility that B^* buyers and S^* sellers enter, with $S^* < B^* \leq \overline{B}$. In order for the buyers to have the incentive to enter, we need $V_b^*(S^*, B^*) \geq \epsilon$. At the same time we have

$$V_b^*(S^*, B^*) = \frac{S^*}{B^*}\left(\frac{1-\delta}{2-\delta-\delta/(B^*-S^*+1)}\right)$$

from (6.9) and Proposition 6.4. It follows that

$$V_b^*(S^*, B^*) < \frac{1-\delta}{2-\delta-\delta/(B^*-S^*+1)} \leq \frac{1-\delta}{2-3\delta/2}.$$

Thus for δ close enough to 1 we have $V_b^*(S^*, B^*) < \epsilon$. Hence there is no equilibrium in which $S^* < B^* \leq \overline{B}$. The other possibility is that $0 < B^* \leq S^*$. In this case we have $p^*(S^*, B^*) \leq 1/2$ from Proposition 6.4, so that $V_b^*(S^*, B^*) = 1 - p^*(S^*, B^*) \geq 1/2 > \epsilon$ (since every buyer is matched immediately when $B^* \leq S^*$). But this implies that $B^* = \overline{B}$, contradicting $B^* \leq S^*$.

We have shown that under the assumption that each agent takes the current value of participating in the market as given when making his entry decision, the only market equilibrium when δ is close to one is one in which no agents enter the market.

An alternative assumption is that each agent anticipates the impact of his entry into the market on the equilibrium price. As in the previous case, if $S^* < B^* \leq \overline{B}$ then the market equilibrium price is close to one when δ is close to one, so that a buyer is better off staying out of the market and avoiding the cost ϵ of entry. Thus there is no equilibrium of this type. If $B^* < S^*$ then the market equilibrium price is less than $1/2$, and even after the entry of an additional buyer it is still at most $1/2$. Thus any buyer not in the market wishes to enter; since $\overline{B} > \overline{S} \geq S^*$ such buyers always exist. Thus there is no equilibrium of this type either. The remaining possibility is that $B^* = S^*$. We shall show that for *every* integer E with $0 \leq E \leq \overline{S}$ there is a market equilibrium of this type, with $S^* = B^* = E$. In such an equilibrium the price is $1/2$, so that no agent prefers to stay out and avoid the entry cost. Suppose that a new buyer enters the market. Then by Proposition 6.4 the price is driven up to $(2-\delta)/(4-3\delta)$ (which is close to 1 when δ is close to 1). The probability of the new buyer being matched with a seller is less than one (it is $S/(S+1)$, since there is now one more buyer than seller), so that the buyer's expected utility is less than $1-(2-\delta)/(4-3\delta) = 2(1-\delta)/(4-3\delta)$. Thus as long as δ is close enough to one that $2(1-\delta)/(4-3\delta)$ is less than ϵ, a buyer not in the market prefers to stay out. Similarly the entry of a new seller will drive the price down close to zero, so that as long as δ is close enough to one a new seller prefers not to enter the market.

Thus when we allow market entry in Model B and assume that each agent fully anticipates the effect of his entry on the market price, there is a multitude of equilibria when $1-\delta$ is small relative to ϵ. In this case, the model predicts only that the numbers of buyers and sellers are the same and that the price is $1/2$.

6.7 A Comparison of the Competitive Equilibrium with the Market Equilibria in Models A and B

The market we have studied initially contains B_0 buyers, each of whom has a "reservation price" of one for one unit of a good, and $S_0 < B_0$ sellers, each of whom has a "reservation price" of zero for the one indivisible unit of the good that she owns. A naïve application of the theory of competitive equilibrium to this market uses the diagram in Figure 6.1. The demand curve D gives the total quantity of the good that the buyers in the market wish to purchase at each fixed price; the supply curve S gives the total quantity the sellers wish to supply to the market at each fixed price. The competitive price is one, determined by the intersection of the curves.

Some, but not all of the models we have studied in this chapter give rise to the competitive equilibrium price of one. Model A (see Section 6.3), in

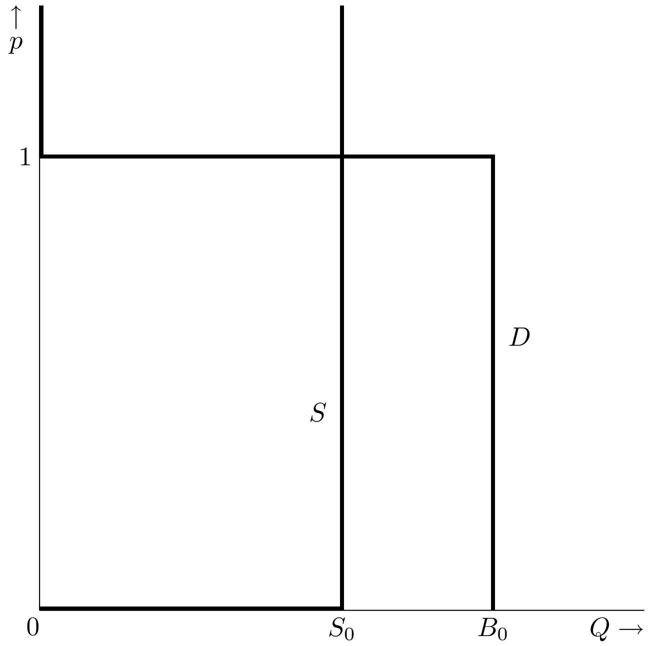


Figure 6.1 Demand and supply curves for the market in this chapter.

which the numbers of buyers and sellers in the market are constant over time, yields an outcome different from the competitive one, even when the discount factor is close to one, if we apply the demand and supply curves to the stocks of traders in the market. In this case the competitive model predicts a price of one if buyers outnumber sellers, and a price of zero if sellers outnumber buyers. However, if we apply the supply and demand curves to the flow of new entrants into the market, the outcome predicted by the competitive model is different. In each period the same number of traders of each type enter the market, leading to supply and demand curves that intersect at all prices from zero to one. Thus under this map of the primitives of the model into the supply and demand framework, the competitive model yields no determinate solution; it includes the price predicted by our market equilibrium, but it also includes every other price between zero and one.

When we add an entry stage to Model A we find that a market equilibrium price of one emerges. In a nondegenerate steady state equilibrium

of a market in which the number of agents is determined endogenously by the agents' entry decisions, the equilibrium price approaches one as the frictions in the market go to zero. This is the "competitive" price when we apply the supply–demand analysis to the numbers of sellers and buyers considering entering the market.

In Model B the unique market equilibrium gives rise to the "competitive" price of one. However, when we start with a pool of agents, each of whom decides whether or not to enter the market, the equilibria no longer correspond to those given by supply–demand analysis. The outcome is sensitive to the way we model the entry decision. If each agent assumes that his own entry into the market will have no effect on the market outcome, then the only equilibrium is that in which no agent enters. If each agent correctly anticipates the impact of his entry on the outcome, then there is a multitude of equilibria, in which equal numbers of buyers and sellers enter. Notice that an equilibrium in which E sellers and buyers enter Pareto dominates an equilibrium in which fewer than E agents of each type enter. This model is perhaps the simplest one in which a coordination problem leads to equilibria that are Pareto dominated.

Notes

Early models of decentralized trade in which matching and bargaining are at the forefront are contained in Butters (1977), Diamond and Maskin (1979), Diamond (1981), and Mortensen (1982a, 1982b). The models in this chapter are similar in spirit to those of Diamond and Mortensen.

Much of the material in this chapter is related to that in the introductory paper Rubinstein (1989). The main difference between the analysis here and in that paper concerns the model of bargaining. Rubinstein (1989) uses a simple strategic model, while here we adopt Nash's axiomatic model. The importance of the distinction between flows and stocks in models of decentralized trade, and the effect of adding an entry decision to such a model was recognized by Gale (see, in particular, (1987)). Sections 6.3, 6.4, and 6.6 include simplified versions of Gale's arguments, as well as ideas developed in the work of Rubinstein and Wolinsky (see, for example, (1985)). A model related to that of Section 6.4 is analyzed in Binmore and Herrero (1988a).

CHAPTER 7

Strategic Bargaining in a Steady State Market

7.1 Introduction

In this chapter and the next we further study the two basic models of decentralized trade that we introduced in the previous chapter (see Sections 6.3 and 6.4). We depart from the earlier analysis by using a simple strategic model of bargaining (like that described in Chapter 3), rather than the Nash bargaining solution, to determine the outcome of each encounter between a buyer and a seller.

The use of a sequential model of bargaining is advantageous in several respects. First, an agent who participates in negotiations that may extend over several periods should consider the possibility either that his partner will abandon him or that he himself will find an alternative partner. It is illuminating to build an explicit model of these strategic considerations. Second, as we saw in the previous chapter, the choice of a disagreement point is not always clear. By using a sequential model, rather than the Nash solution, we avoid the need to specify an exogenous disagreement point. Finally, although the model we analyze here is relatively simple, it supplies a framework for analyzing more complex markets. The strategic approach lends itself to variations in which richer economic institutions can be modeled.

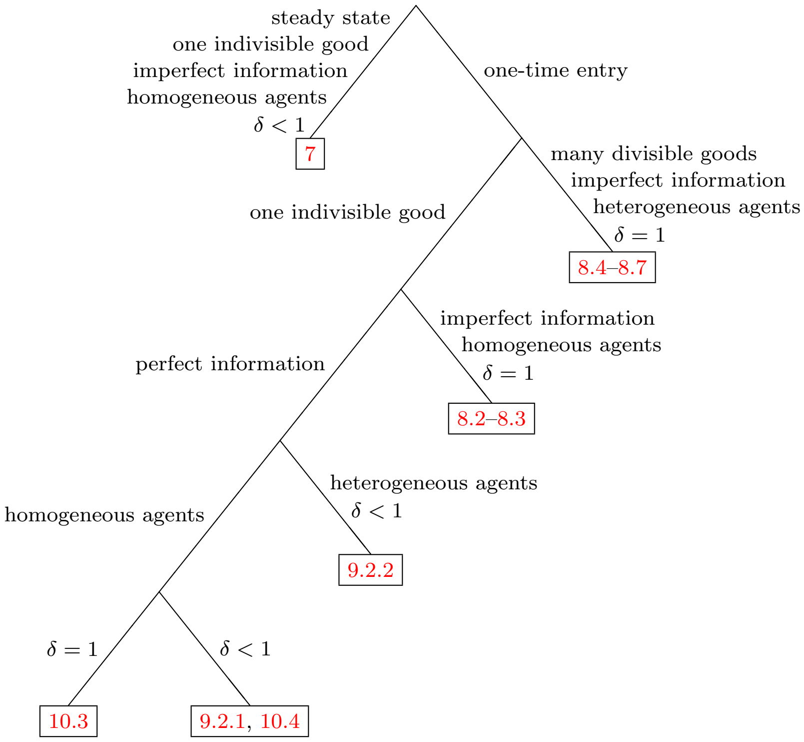


Figure 7.1 Strategic models of markets with random matching. The figure should be read from the top down. The numbers in boxes are the chapters and sections in which models using the indicated assumptions are discussed. Thus, for example, a model with one-time entry, one indivisible good, imperfect information, homogeneous agents, and $\delta = 1$ is discussed in Sections 8.2 and 8.3.

The models that we study in this and the following chapters differ in the assumptions they make about the evolution of the number of participants in the market, the nature of the goods being traded, the information possessed by the agents, and the agents' preferences. The various combinations of assumptions that we investigate in markets with random matching are summarized in Figure 7.1.

7.2 The Model

The model we study here has the structure of Model A of the previous chapter (see Section 6.3), with a single exception: the outcome of bargain-

ing is determined by a simple strategic model rather than being given by Nash's bargaining solution.

Goods There is a single indivisible good which is traded for some quantity of a divisible good ("money").

Time Time is discrete and is indexed by the integers; it stretches infinitely in both directions.

Economic Agents The economic agents are (potential) buyers and sellers. Each seller enters the market with one unit of the indivisible good; each buyer enters with one unit of money. Each agent is concerned about the agreement price p and the period t in which agreement is reached. Each agent's preferences on lotteries over pairs (p, t) satisfy the assumptions of von Neumann and Morgenstern. Each seller's preferences are represented by the utility function $\delta^t p$, while each buyer's preferences are represented by $\delta^t(1-p)$. The common discount factor δ satisfies $0 < \delta < 1$. If an agent never trades, then he obtains the utility of zero.

Bargaining During the first phase of a period the members of any matched pair bargain. A random device selects one of the agents to propose a price, which the other agent may accept or reject. The probability that a particular agent is chosen to propose a price is $1/2$, independent of all past events. In the event of acceptance, a transaction is completed at the agreed-upon price, and the two agents leave the market. In the event of rejection, the two agents participate in the matching process.

Matching In the matching phase each agent in the market (whether or not he had a partner at the beginning of the period) is matched, with positive probability, with an agent of the opposite type. Each seller is matched with a new buyer with probability α, and each buyer is matched with a new seller with probability β. These events are independent of each other and of all past events (including whether or not the agent had a partner at the beginning of the period), and the probabilities are constant over time. Each agent who is matched anew *must* abandon his old partner (if any); it is not possible to have more than one partner simultaneously. Thus an agent who has a partner at the beginning of period t and fails to reach agreement in the bargaining phase continues negotiating in period $t+1$ with this partner if neither of them is newly matched; he starts new negotiations if he is newly matched, and does not participate in bargaining in period $t+1$ if his partner is newly matched and he is not. If a buyer

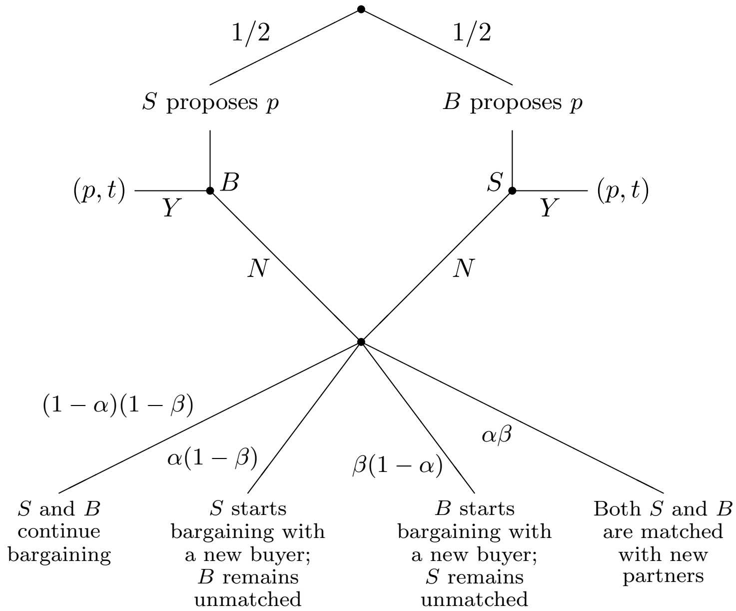


Figure 7.2 The structure of events within some period t. S and B stand for the seller and the buyer, and Y and N stand for acceptance and rejection. The numbers beside the branches are the probabilities with which the branches occur.

and seller are partners at the beginning of period t, there are thus four possibilities for period $t+1$. With probability $(1-\alpha)(1-\beta)$ the pair continues bargaining; with probability $\alpha(1-\beta)$ the seller starts bargaining with a new partner, while the buyer is idle; with probability $\beta(1-\alpha)$ the buyer starts bargaining with a new partner, while the seller is idle; and with probability $\alpha\beta$ both traders start bargaining with new partners.

The structure of events within some period t is illustrated in Figure 7.2.

In the bargaining game of alternating offers studied in Chapter 3, each player is under pressure to reach an agreement because he is impatient. Here also each agent is impatient. But there is an additional pressure: the risk that his partner will be matched anew. Each agent is thus concerned about his partner's probability of being matched with another agent.

Note that when a trader is matched with a new partner, he does *not* have the option of continuing negotiation with his current partner. At every new encounter an agent is constrained to abandon his current partner. However,

although in this model an agent does not decide whether to abandon his partner (cf. the models in Sections 3.12 and 9.4), in equilibrium this act of abandonment does not conflict with optimization.

Note also that the model is not formally a game, since we have not specified the set of players. The primitives of the model are the (constant) probabilities α and β of agents being matched with new partners and not either the sets or the numbers of sellers and buyers in the market. These assumptions are appropriate in a large market in which the variations are small. In such a case an agent may ignore information about the names of his partners and the exact numbers of sellers and buyers, and base his behavior merely on his evaluation of the speed with which he finds potential partners and the intensity of his fear that his partner will abandon him. A model in which the sets of sellers and buyers are the primitives is studied in Section 8.2.

We can link a model in which the primitives are the numbers S of sellers and B of buyers with the current model by adding details about the matching technology. We can assume, for example, that the probability of being matched depends only on these numbers, that these numbers are constant over time, and that there are M new matches in each period. If the numbers S and B are large (so that the probability of a given agent being rematched with his current partner is small), then this technology gives approximately $\alpha = M/S$ and $\beta = M/B$.

7.3 Market Equilibrium

When a seller and a buyer are matched they start a bargaining game in which, in each period that they remain matched, one of them is selected to make an offer. The bargaining stops either when an agreement is reached or when at least one of the parties is matched with a new partner. Thus, the history of negotiation in a particular match is a sequence of selections of a proposer, offers, and reactions. After a history that ends with the selection of a proposer, that agent makes an offer; after a history that ends with an offer by an agent, the other has to respond.

We define an agent's *strategy* to be a function that assigns to every possible history of events within a match either a price or a response (Y or N), according to whether the agent has to make an offer or to respond to an offer. Thus an agent's strategy has the same structure as that of a strategy in a bargaining game of alternating offers in which the proposer is chosen randomly each period (see the end of Section 3.10.3). By defining a strategy in this way, we are assuming that each agent uses the same rule of behavior in every bargaining encounter. We refer to this assumption as *semi-stationarity*. Behind the definition lies the assumption that

each agent can recall perfectly the events in any particular bargaining encounter and may respond differently to different histories while bargaining with the same agent. However, he cannot condition his behavior on the events that occurred in any period before he started bargaining with his current partner. In particular, when he is matched with a new partner he does not know whether he was ever matched with that partner in the past. However, an agent continues to recognize his partner until the match dissolves.

We make the assumption that the behavior of each agent in a match is independent of the events in previous matches in which he participated in order to simplify the analysis. However, note that this would not be a natural assumption were we to consider a model in which the agents are asymmetrically informed, since in this case each agent gathers information while bargaining.

We restrict attention to the case in which each agent of a given type (seller, buyer) uses the *same* (semi-stationary) strategy. Given a pair of strategies—one for every seller and one for every buyer—and the probabilities of matches, we can calculate the expected utilities of matched and unmatched sellers and buyers at the beginning of a period, discounted to that period. Because each agent's behavior is semi-stationary, these utilities are independent of the period. Let V_s be the expected utility of an unmatched seller and let V_b be the expected utility of an unmatched buyer. Let W_s and W_b be the corresponding expected utilities for matched sellers and buyers. (Note that these expected utilities, in contrast to the ones denoted V_s and V_b in Chapter 6, are calculated *after* the matching process.) The variables V_s, V_b, W_s, and W_b are functions of the pair of the strategies and satisfy the following conditions.

$$V_s = \delta[\alpha W_s + (1-\alpha)V_s] \tag{7.1}$$

$$V_b = \delta[\beta W_b + (1-\beta)V_b]. \tag{7.2}$$

A pair that is bargaining in period t continues to do so in period $t+1$ if and only if neither is matched with a new partner, an event with probability $(1-\alpha)(1-\beta)$. Thus the probability of breakdown is $q = 1-(1-\alpha)(1-\beta)$. Conditional on at least one of the agents being matched with a new partner, the seller is matched with a new buyer with probability α/q and remains unmatched with probability $\beta(1-\alpha)/q$. Thus the seller's payoff in the event of breakdown in period t (discounted to period t) is

$$U_s = \delta[\alpha W_s + \beta(1-\alpha)V_s]/q. \tag{7.3}$$

Similarly, the expected utility of a buyer in the event of breakdown is

$$U_b = \delta[\beta W_b + \alpha(1-\beta)V_b]/q. \tag{7.4}$$

The game between the members of a matched pair is very similar to the game analyzed in Section 4.2. It depends on the expected utilities of the agents in the event of breakdown (the values of which are determined within the model). When these utilities are u_s and u_b its structure is as follows. At the start of each period one of the players is selected, with probability $1/2$, to propose a price in $[0, 1]$. The other player responds by accepting or rejecting the proposal. In the event of acceptance, the game ends; in the event of rejection, there is a chance move which terminates the game with probability q, giving the players the payoffs (u_s, u_b). With probability $1 - q$ the game continues: play passes to the next period. We denote this game $\Gamma(u_s, u_b)$. (Notice the differences between this game and the game analyzed in Section 4.2: the proposer is chosen randomly at the start of every period, and the outcome in the event of breakdown is not necessarily the worst outcome in the game.)

Recall that V_s, V_b, W_s, and W_b, and hence U_s and U_b, are functions of the pair of strategies; for clarity we now record this dependence in the notation.

Definition 7.1 A *market equilibrium* is a pair (σ^*, τ^*) of (semi-stationary) strategies that is a subgame perfect equilibrium of the game $\Gamma(u_s, u_b)$, where $u_s = U_s(\sigma^*, \tau^*)$ and $u_b = U_b(\sigma^*, \tau^*)$.

7.4 Analysis of Market Equilibrium

We now characterize market equilibrium.

Proposition 7.2 *There is a unique market equilibrium. In this equilibrium the seller always proposes the price x^* and accepts any price at least equal to y^*, and the buyer always proposes the price y^* and accepts any price at most equal to x^*, where*

$$x^* = \frac{2(1-\delta) + \delta\alpha - \delta(1-\delta)(1-\alpha)(1-\beta)}{2(1-\delta) + \delta\alpha + \delta\beta} \tag{7.5}$$

$$y^* = \frac{\delta\alpha + \delta(1-\delta)(1-\alpha)(1-\beta)}{2(1-\delta) + \delta\alpha + \delta\beta}. \tag{7.6}$$

Proof. First, using the methods of Section 4.2 we can show that for any given pair of numbers (u_s, u_b) for which $u_s + u_b < 1$, the game $\Gamma(u_s, u_b)$ has a unique subgame perfect equilibrium, which is characterized by a pair of numbers (x^*, y^*). In this equilibrium, the seller always offers the price x^*, and accepts any price at least equal to y^*; the buyer always offers y^*, and accepts any price at most equal to x^*. The pair (x^*, y^*) is the solution

of the following pair of equations.

$$y^* = qu_s + (1-q)\delta(x^* + y^*)/2 \tag{7.7}$$
$$1 - x^* = qu_b + (1-q)\delta(1 - x^*/2 - y^*/2) \tag{7.8}$$

The payoffs in this equilibrium are $(x^* + y^*)/2$ for the seller, and $1 - (x^* + y^*)/2$ for the buyer.

Next, we verify that in every market equilibrium (σ^*, τ^*) we have $U_s + U_b < 1$. From (7.1) we have $V_s < W_s$; from (7.3) it follows that $U_s < W_s$. Similarly $U_b < W_b$, so that $U_s + U_b < W_s + W_b$. Since $W_s + W_b$ is the expectation of a random variable all values of which are at most equal to the unit surplus available, we have $W_s + W_b \leq 1$.

Thus a market equilibrium strategy pair has to be such that the induced variables V_s, V_b, W_s, W_b, U_s, U_b, x^*, and y^* satisfy the four equations (7.1), (7.2), (7.3), (7.4), the two equations (7.7) and (7.8) with $u_s = U_s$ and $u_b = U_b$, and the following additional two equations.

$$W_s = (x^* + y^*)/2 \tag{7.9}$$
$$W_b = 1 - (x^* + y^*)/2. \tag{7.10}$$

It is straightforward to verify that solution to these equations, which is unique, is that given in (7.5) and (7.6). □

So far we have restricted agents to use semi-stationary strategies: each agent is constrained to behave the same way in every match. We now show that if every buyer uses the (semi-stationary) equilibrium strategy described above, then any given seller cannot do better by using different bargaining tactics in different matches. A symmetric argument applies to sellers. In other words, the equilibrium we have found remains an equilibrium if we extend the set of strategies to include behavior that is not semi-stationary.

Consider some seller. Suppose that every buyer in the market is using the equilibrium strategy described above, in which he always offers y^* and accepts no price above x^* whenever he is matched. Suppose that the seller can condition her actions on her entire history in the market. We claim that the strategy of always offering x^* and accepting no price below y^* is optimal among *all* possible strategies.

The environment the seller faces after any history can be characterized by the following four states:

e_1: the seller has no partner

e_2: the seller has a partner, and she has been chosen to make an offer

e_3: the seller has a partner, and she has to respond to the offer y^*

e_4: agreement has been reached

Each agent's history in the market corresponds to a sequence of states. The initial state is e_1. A strategy of the seller can be characterized as a function that assigns to each sequence of states an action of either *stop* or *continue*. The only states in which the action has any effect are e_2 and e_3. In state e_2, a buyer will accept any offer at most equal to x^*, so that any such offer stops the game. However, given the acceptance rule of each buyer, it is clearly never optimal for the seller to offer a price less than x^*. Thus *stop* in e_2 means make an offer of x^*, while *continue* means make an offer in excess of x^*. In state e_3, *stop* means accept the offer y^*, while *continue* means reject the offer.

The actions of the seller determine the probabilistic transitions between states. Independent of the seller's action the system moves from state e_1 to states e_2 and e_3, each with probability $\alpha/2$, and remains in state e_1 (the seller remains unmatched) with probability $1-\alpha$. (In this case the action *stop* does not stop the game.) State e_4 is absorbing: once it is reached, the system remains there. The transitions from states e_2 and e_3 depend on the action the seller takes. If the seller chooses *stop*, then in either case the system moves to state e_4 with probability one. If the seller chooses *continue*, then in either case the system moves to the states e_1, e_2, and e_3 with probabilities $(1-\alpha)\beta$, $[1-(1-\alpha)\beta]/2$, and $[1-(1-\alpha)\beta]/2$, respectively.

To summarize, the transition matrix when the seller chooses *stop* is

$$\begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{array} \begin{array}{c} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} \\ \begin{pmatrix} 1-\alpha & \alpha/2 & \alpha/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

and that when the seller chooses *continue* is

$$\begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{array} \begin{array}{c} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} \\ \begin{pmatrix} 1-\alpha & \alpha/2 & \alpha/2 & 0 \\ (1-\alpha)\beta & [1-(1-\alpha)\beta]/2 & [1-(1-\alpha)\beta]/2 & 0 \\ (1-\alpha)\beta & [1-(1-\alpha)\beta]/2 & [1-(1-\alpha)\beta]/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

The seller gets a payoff of zero unless she chooses *stop* at one of the states e_2 or e_3. If she chooses *stop* in state e_2, then her payoff is x^*, while if she chooses *stop* in e_3 then her payoff is y^*.

This argument shows that the seller faces a Markovian decision problem. Such a problem has a stationary solution (see, for example, Derman (1970)). That is, there is a subset of states with the property that it is optimal for the seller to choose *stop* whenever a state in the subset is reached. Choosing *stop* in either e_1 or e_4 has no effect on the evolution of the system, so we can restrict attention to rules that choose *stop* in some subset of $\{e_2, e_3\}$. If this subset is empty (*stop* is never chosen), then the payoff is zero; since the payoff is otherwise positive, an optimal stopping set is never empty. Now suppose that *stop* is chosen in e_3. If *stop* is also chosen in e_2, the seller receives a payoff of x^*, while if *continue* is chosen in e_2, the best that can happen is that e_3 is reached in the next period, in which case the seller receives a payoff of y^*. Since $y^* < x^*$, it follows that it is better to choose *stop* than *continue* in e_2 if *stop* is chosen in e_3. Thus the remaining candidates for an optimal stopping set are $\{e_2\}$ and $\{e_2, e_3\}$. A calculation shows that the expected utilities of these stopping rules are the same, equal to[1] $\alpha/[2(1-\delta)+\delta\alpha+\delta\beta]$. Thus $\{e_2, e_3\}$ is an optimal stopping set: it is optimal for a seller to use the semi-stationary strategy described in Proposition 7.2 even when she is not restricted to use a semi-stationary strategy. A similar argument applies to the buyer's strategy.

Finally, we note that although an agent who is matched with a new partner is forced to abandon his current partner, this does not conflict with optimal behavior in equilibrium. Agreement is reached immediately in every match, so that giving an agent the option of staying with his current partner has no effect, given the strategies of all other agents.

7.5 Market Equilibrium and Competitive Equilibrium

The fact that the discount factor δ is less than 1 creates a friction in the market—a friction that is absent from the standard model of a competitive market. If we wish to compare the outcome with that predicted by a competitive analysis, we need to consider the limit of the market equilibrium as δ converges to 1.

One limit in which we may be interested is that in which δ converges to 1 while α and β are held constant. From (7.5) and (7.6) we have

$$\lim_{\delta\to 1} x^* = \lim_{\delta\to 1} y^* = \frac{\alpha}{\alpha+\beta}.$$

Thus in the limit the surplus is divided in proportion to the matching probabilities. This is the same as the result we obtained in Model A of

[1] Consider, for example, the case in which the stopping set is $\{e_2\}$. Let E be the expected utility of the seller. Then $E = (1-\alpha)\delta E + (\alpha/2)x^* + (\alpha/2)y^*$, which yields the result.

Chapter 6 (see Section 6.3), where we used the Nash bargaining solution, rather than a strategic model, to analyze a market.

This formula includes the probabilities that a seller and buyer are matched with a partner in any given period, but not the numbers of sellers and buyers in the market. In order to compare the market equilibrium with the equilibrium of a competitive market, we need to relate the probabilities to the population size. Suppose that the probabilities are derived from a matching technology in a model in which the primitives are the numbers S and B of sellers and buyers in the market. Specifically assume that $\alpha = M/S$ and $\beta = M/B$ for some fixed M, interpreted as the number of matches per unit of time. Then the limit of the market equilibrium price as δ converges to 1 is $B/(S + B)$.

Now suppose that the number of buyers in the market exceeds the number of sellers. Then the competitive equilibrium, applied to the supply–demand data for the agents in the market, yields a competitive price of one (cf. the discussion in Section 6.7). By contrast, the model here yields an equilibrium price strictly less than one. Note, however, that if we apply the supply–demand analysis to the flows of agents into the market, then every price equates demand and supply, so that a market equilibrium price is a competitive price (see Section 6.7).

If we generalize the model of this chapter to allow the agents' reservation prices to take an arbitrary finite number of different values, then the demand and supply curves of the stocks of buyers and sellers in the market in each period are step functions. Suppose, in this case, that the probability of an individual being matched with an agent of a particular type is proportional to the number of agents of that type in the market. Then the limit of the unique market equilibrium price p^* as $\delta \to 1$ has the property that the area above the horizontal line at p^* and below the demand curve is equal to the area below this horizontal line and above the supply curve (see Figure 7.3). That is, the limiting market equilibrium price equates the demand and supply "surpluses". (See Gale (1987, Proposition 11).) Note that for the special case in which there are S identical sellers with reservation price 0 and $B > S$ identical buyers with reservation price 1, the limiting market equilibrium price p^* given by this condition is precisely $B/(S + B)$, as we found above.

Notes

The main model and result of this chapter are due to Rubinstein and Wolinsky (1985). The extension to markets in which the supply and demand functions are arbitrary step-functions (discussed at the end of the last section) is due to Gale (1987, Section 6).

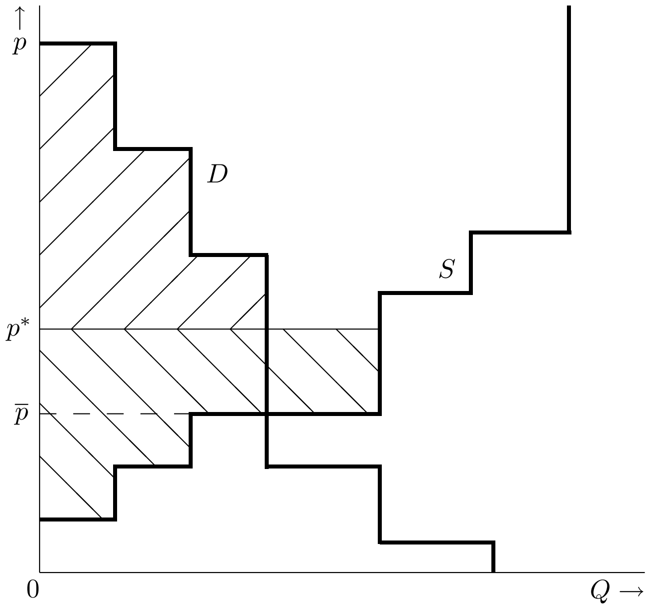


Figure 7.3 Demand and supply curves in a market of heterogeneous agents. The heavy lines labeled D and S are the demand and supply curves of the stocks of buyers and sellers in the market in each period. In this market the agents have a finite number of different reservation prices. The market equilibrium price p^* has the property that the shaded areas are equal. The price that equates supply and demand is $\overline{p}$.

Binmore and Herrero (1988b) investigate the model under the assumption that agents' actions in bargaining encounters are not independent of their personal histories. If agents' strategies are not semi-stationary then an agent who does not know his opponent's personal history cannot figure out how the opponent will behave in the bargaining encounter. Therefore, the analysis of this chapter cannot be applied in a straightforward way; Binmore and Herrero introduce a new solution concept (which they call "security equilibrium"). Wolinsky (1987) studies a model in which each agent chooses the intensity with which he searches for an alternative partner. Wolinsky (1988) analyzes the case in which transactions are made by an auction, rather than by matching and bargaining. In the models in all these papers the agents are symmetrically informed. Wolinsky (1990) initiates the investigation of models in which agents are asymmetrically informed (see also Samuelson (1992) and Green (1992)).

Models of decentralized trade that explicitly specify the process of trade are promising vehicles for analyzing the role and value of money in a market. Gale (1986d) studies a model in which different agents are initially endowed with different divisible goods and money, and all transactions must be done in exchange for money. He finds that there is a great multiplicity of inefficient equilibria. Kiyotaki and Wright (1989) study a model in which each agent is endowed with one unit of one of the several indivisible goods in the market, and there is only one possible exchange upon which a matched pair can agree. In some equilibria of the model some goods play the role of money: they are traded simply as a medium of exchange.

CHAPTER 8

Strategic Bargaining in a Market with One-Time Entry

8.1 Introduction

In this chapter we study two strategic models of decentralized trade in a market in which all potential traders are present initially (cf. Model B of Chapter 6). In the first model there is a single indivisible good that is traded for a divisible good ("money"); a trader leaves the market once he has completed a transaction. In the second model there are many divisible goods; agents can make a number of trades before departing from the market. (This second model is close to the standard economic model of competitive markets.)

We focus on the conditions under which the outcome of decentralized trade is competitive; we point to the elements of the models that are critical for a competitive outcome to emerge. In the course of the analysis, several issues arise concerning the nature of the information possessed by the agents. In Chapter 10 we return to the first model and study in detail the role of the informational assumptions in leading to a competitive outcome.

8.2 A Market in Which There Is a Single Indivisible Good

The first model is possibly the simplest model that combines pairwise meetings with strategic bargaining.

Goods A single indivisible good is traded for some quantity of a divisible good ("money").

Time Time is discrete and is indexed by the nonnegative integers.

Economic Agents In period 0, S identical sellers enter the market with one unit of the indivisible good each, and $B > S$ identical buyers enter with one unit of money each. No more agents enter at any later date. Each individual's preferences on lotteries over the price p at which a transaction is concluded satisfy the assumptions of von Neumann and Morgenstern. Each seller's preferences are represented by the utility function p, and each buyer's preferences are represented by the utility function $1 - p$ (i.e. the reservation values of the seller and buyer are zero and one respectively, and no agent is impatient). If an agent never trades then his utility is zero.

Matching In each period any remaining sellers and buyers are matched pairwise. The matching technology is such that each seller meets exactly one buyer and no buyer meets more than one seller in any period. Since there are fewer sellers than buyers, $B - S$ buyers are thus left unmatched in each period. The matching process is random: in each period all possible matches are equally probable, and the matching is independent across periods.

Although this matching technology is very special, the result below can be extended to other technologies in which the probabilities of any particular match are independent of history.

Bargaining After a buyer and a seller have been matched they engage in a short bargaining process. First, one of the matched agents is selected randomly (with probability 1/2) to propose a price between 0 and 1. Then the other agent responds by accepting the proposed price or rejecting it. Rejection dissolves the match, in which case the agents proceed to the next matching stage. If the proposal is accepted, the parties implement it and depart from the market.

Information We assume that the agents have information only about the index of the period and the names of the sellers and buyers in the market. (Thus they know more than just the *numbers* of sellers and buyers in the market.) When matched, an agent recognizes the name

of his opponent. However, agents do not remember the past events in their lives. This may be because their memories are poor or because they believe that their personal experiences are irrelevant. Nor do agents receive any information about the events in matches in which they did not take part.

These assumptions specify an extensive game. Note that since the agents forget their own past actions, the game is one of "imperfect recall". We comment briefly on the consequences of this at the end of the next section.

8.3 Market Equilibrium

Given our assumption about the structure of information, a *strategy* for an agent in the game specifies an offer and a response function, possibly depending on the index of the period, the sets of sellers and buyers still in the market, and the name of the agent's opponent. To describe a strategy precisely, note that there are two circumstances in which agent i has to move. The first is when the agent is matched and has been selected to make an offer. Such a situation is characterized by a triple (t, A, j), where t is a period, A is a set of agents that includes i (the set of agents in the market in period t), and j is a member of A of the opposite type to i (i's partner). The second is when the agent has to respond to an offer. Such a situation is characterized by a four-tuple (t, A, j, p), where t is a period, A is a set of agents that includes i, j is a member of A of the opposite type to i, and p is a price in $[0, 1]$ (an offer by j). Thus a strategy for agent i is a pair of functions, the first of which associates a price in the interval $[0, 1]$ with every triple (t, A, j), and the second of which associates a member of the set $\{Y, N\}$ ("accept", "reject") with every four-tuple (t, A, j, p).

The spirit of the solution concept we employ is close to that of sequential equilibrium. An agent's strategy is required to be optimal not only at the beginning of the game but also at every other point at which the agent has to make a decision. A strategy induces a plan of action starting at any point in the game. We now explain how each agent calculates the expected utility of each such plan of action.

First, suppose that agent i is matched and has been selected to make an offer. In such a situation i's information consists of (t, A, j), as described above. The behavior of every other agent in A depends only on t, A, and the agent with whom that agent is matched (if any). Thus the fact that i does not know the events that have occurred in the past is irrelevant, because neither does any other agent, so that no other agent's actions are conditioned on these events. In this case, agent i's information is sufficient, given the strategies of the other agents, to calculate the moves of his future

partners, and thus find the expected utility of any plan of action starting at t.

Second, suppose that agent i has to respond to an offer. In this case i's information consists of a four-tuple (t, A, j, p), as described above. If he accepts the offer then his utility is determined by p. If he rejects the offer, then his expected utility is determined by the events in other matches (which determine the probabilities with which he will be matched with any remaining agents) and the other agents' strategies. If p is the offer that is made when all agents follow their equilibrium strategies, then the agent uses these strategies to form a belief about the events in other matches. If p is different from the offer made in the equilibrium—if the play of the game has moved "off the equilibrium path"—then the notion of sequential equilibrium allows the agent some freedom in forming his belief about the events in other matches. We assume that the agent believes that the behavior of all agents in any simultaneous matches, and in the future, is still given by the equilibrium strategies. Even though he has observed an action that indicates that some agent has deviated from the equilibrium, he assumes that there will be no further deviations. Given that the agent expects the other agents to act in the future as they would in equilibrium, he can calculate his expected utility from each possible plan of action starting at that point.

Definition 8.1 A *market equilibrium* is a strategy profile (a strategy for each of the $S + B$ agents), such that each agent's strategy is optimal at every point at which the agent has to make a choice, on the assumption that all the actions of the other agents that he does not observe conform with their equilibrium strategies.

Proposition 8.2 *There exists a market equilibrium, and in every such equilibrium every seller's good is sold at the price of one.*

This result has two interesting features. First, although we do not assume that all transactions take place at the same price, we obtain this as a result. Second, the equilibrium price is the competitive price.

Proof of Proposition 8.2. We first exhibit a market equilibrium in which all units of the good are sold at the price of one. In every event all agents offer the price one, every seller accepts only the price one, and every buyer accepts any price. The outcome is that all goods are transferred, at the price of one, to the buyers who are matched with sellers in the first period. No agent can increase his utility by adopting a different strategy. Suppose, for example, that a seller is confronted with the offer of a price less than one (an event inconsistent with equilibrium). If she rejects this offer, then she

will certainly be matched in the next period. Under our assumption that she believes, despite the previous inconsistency with equilibrium, that all agents will behave in the future according to their equilibrium strategies, she believes that she will sell her unit at the price one in the next period. Thus it is optimal for her to reject the offer.

We now prove that there is no other market equilibrium outcome. We use induction on the number of sellers in the market. First consider the case of a market with a single seller ($S = 1$). In this case the set of agents in the market remains the same as long as the market continues to operate. Thus if no transaction has taken place prior to period t, then at the beginning of period t, before a match is established, the expected utilities of the agents depend only on t. For any given strategy profile let $V_i^b(t)$ and $V^s(t)$ be these expected utilities of buyer i and the seller, respectively.

Let m be the infimum of $V^s(t)$ over all market equilibria and all t. Fix a market equilibrium. Since there is just one unit of the good available in the economy, we have $\sum_{i=1}^{B} V_i^b(t) \leq 1 - m$ for all t. Thus for each t there is a buyer for whom $V_i^b(t+1) \leq (1-m)/B$. Suppose the seller adopts the strategy of proposing the price $1 - \epsilon - (1-m)/B$, and rejecting all lower prices, for some $\epsilon > 0$. Eventually she will meet, say in period t, a buyer for whom $V_i^b(t+1) \leq (1-m)/B$. The optimality of this buyer's strategy demands that he accept this offer, so that the seller obtains a utility of $1 - \epsilon - (1-m)/B$. Thus $V^s(t) \geq 1 - \epsilon - (1-m)/B$. Therefore $m \geq 1 - \epsilon - (1-m)/B$, and hence $m \geq 1 - \epsilon B/(B-1)$ for any $\epsilon > 0$, which means that $m = 1$.

Now assume the proposition is valid if the number of sellers in the markets is strictly less than $\overline{S}$. Fix a set of sellers of size $\overline{S}$. For any given strategy profile let $V_j^s(t)$ and $V_i^b(t)$ be the expected utilities of seller j and buyer i, respectively, at the beginning of period t (before any match is established) if all the $\overline{S}$ sellers in the set and all B buyers remain in the market. We shall show that for all market equilibria in a market containing the $\overline{S}$ sellers and B buyers we have $V_j^s(0) = 1$ for every seller j. Let m be the infimum of $V_j^s(t)$ over all market equilibria, all t, and all j. Fix a market equilibrium. For all t we have $\sum_{i=1}^{B} V_i^b(t) \leq (1-m)\overline{S}$. Therefore, in any period t there exists some buyer i such that $V_i^b(t+1) \leq (1-m)\overline{S}/B$. Consider a seller who adopts the strategy of demanding the price $1-\epsilon-(1-m)\overline{S}/B$ and not agreeing to less as long as the market contains the $\overline{S}$ sellers and B buyers. Either she will be matched in some period t with a buyer for whom $V_i^b(t+1) \leq (1-m)\overline{S}/B$ who will then agree to that price, or some other seller will transact beforehand. In the first case the seller's utility will be $1 - \epsilon - (1-m)\overline{S}/B$, while in the second case it will be 1 by the inductive hypothesis. Since a seller can always adopt this strategy, we have

$V_j^s(t) \geq 1 - \epsilon - (1-m)\overline{S}/B$. Therefore $m \geq 1 - \epsilon - (1-m)\overline{S}/B$, and hence $m \geq 1 - \epsilon B/(B - \overline{S})$ for any $\epsilon > 0$, which means that $m = 1$. □

There are three points to notice about the result. First, it does not state that there is a unique market equilibrium—only that the price at which each unit of the good is sold in every market equilibrium is the same. There are in fact other market equilibria—for example, ones in which all sellers reject all the offers made by a particular buyer. Second, the proof remains unchanged if we assume that agents do not recognize the name of their opponents. The informational assumptions we have made allow us to conclude that, at the beginning of each period, the expected utilities of being in the market depend only on the index of the period. Assuming that agents cannot recognize their opponents does not affect this conclusion. Third, the proof reveals the role played by the surplus of buyers in determining the competitive outcome. The probability that a seller is matched in any period is one, while this probability is less than one for a buyer. Although there is no impatience in the model, the situation is somewhat similar to that of a sequential bargaining game in which the seller's discount factor is 1 and the buyer's discount factor is $S/B < 1$.

As we mentioned above, the model is a game with imperfect recall. Each agent forgets information that he possessed in the past (like the names of agents with whom he was matched and the offers that were made). The only information that an agent recalls is the time and the set of agents remaining in the market. The issue of how to interpret the assumption of imperfect recall is subtle; we do not discuss it in detail (see Rubinstein (1991) for more discussion). We simply remark that the assumption we make here has implications beyond the fact that the behavior of an agent can depend only on time and the set of agents remaining in the market. The components of an agent's strategy that specify his actions after arbitrary histories can be interpreted as reflecting his beliefs about what other agents expect him to do in such cases. Thus our assumption means also that no event in the past leads an agent to change his beliefs about what other agents expect him to do.

8.4 A Market in Which There Are Many Divisible Goods

The main differences between the model we study here and that of the previous two sections are that the market here contains many divisible goods, rather than a single indivisible good, and that agents may make many transactions before departing from the market. We begin with an outline of the model.

There is a continuum of agents in the market, trading m divisible goods. Time is discrete and is indexed by the nonnegative integers. All agents enter the market simultaneously in period 0; each brings with him a bundle of goods, which may be stored costlessly. In period 0 and all subsequent periods there is a positive probability that any given agent is matched with a trading partner. Once a match is formed, one of the parties is selected at random to propose a trade (an exchange of goods). The other agent may accept or reject this proposal. If he rejects it then he may, if he wishes, leave the market. Agents who remain in the market are matched anew with positive probability each period and may execute a sequence of transactions. All matches cease after one period: even if an agent who is matched in period t is not matched with a new partner in period $t+1$, he must abandon his old partner. An agent obtains utility from the bundle he holds when he leaves the market. Note that agents may not leave the market immediately after accepting an offer; they may leave *only* after rejecting an offer. Although this assumption lacks intuitive appeal, it formalizes the idea that an agent who is about to depart from the market always has a "last chance" to receive an offer.

We now spell out the details of the model.

Goods There are m divisible goods; a bundle of goods is a member of $\mathcal{R}^m_+$.

Time Time is discrete and is indexed by the nonnegative integers.

Economic Agents There is a continuum of agents in the market. Each agent is characterized by the initial bundle with which he enters the market and his von Neumann–Morgenstern utility function over the union of the set $\mathcal{R}^m_+$ of feasible consumption bundles and the event D of staying in the market forever. Each agent chooses the period in which to consume, and is indifferent about the timing of his consumption (i.e. is not impatient). The agents initially present in the market are of a finite number K of *types*. All members of any given type k have the same utility function $u_k\colon \mathcal{R}^m_+ \cup \{D\} \to \mathcal{R} \cup \{-\infty\}$ and the same initial bundle $\omega_k \in \mathcal{R}^m_+$. For each type k there is initially the measure n_k of agents in the market (with $\sum_{k=1}^K n_k = 1$). Each utility function u_k is restricted as follows. There is a continuous function $\phi_k\colon \mathcal{R}^m_+ \to \mathcal{R}$ that is increasing and strictly concave on the interior of $\mathcal{R}^m_+$ and satisfies $\phi_k(x) = 0$ if x is on the boundary of $\mathcal{R}^m_+$. Let $\overline{\phi} > 0$ be a number, and let $X_k = \{x \in \mathcal{R}^m_+\colon \phi_k(x) \geq \overline{\phi}\}$. Then u_k is given by $u_k(x) = \phi_k(x)$ if $x \in X_k$ and $u_k(x) = -\infty$ for all other x (including $x = D$). (The number $\overline{\phi}$ can be interpreted as the minimal utility necessary to survive. The assumption that $u_k(D) = -\infty$ means that agents must leave the market eventually.) Further, we assume that

$\omega_k \in X_k$. An interpretation of the concavity of the utility functions is that each agent is risk-averse. We make two further assumptions on the utility functions.

1. For each k there is a unique tangent to each indifference curve of u_k at every point in X_k.
2. Fix some type k and some nonzero vector $p \in \mathcal{R}^m_+$. Consider the set $S(k, p)$ of bundles c for which the tangent to the indifference curve of u_k through c is $\{x: px = pc\}$ (i.e. $S(k, p)$ is k's "income-expansion" path at the price vector p). Then for every vector $z \in \mathcal{R}^m$ for which $pz > 0$ there exists a positive integer L such that $u_k(c + z/L) > u_k(c)$ for every c in $S(k, p)$.

The first assumption is weaker than differentiability of u_k on X_k (since it relates only to the indifference curves of u_k). Note that it guarantees that for each vector $z \in \mathcal{R}^m$ and each bundle c in $S(k, p)$ we can find an integer L such that $u_k(c + z/L) > u_k(c)$. The second assumption imposes the stronger condition that for each vector $z \in \mathcal{R}^m$ we can find a single L such that $u_k(c+z/L) > u_k(c)$ for *all* c in $S(k, p)$. This second assumption is illustrated in Figure 8.1. (It is related to Gale's (1986c) assumption that the indifference curves of the utility function have uniformly bounded curvature.)

Matching In every period each agent is matched with a partner with probability $0 < \alpha < 1$ (independent of all past events). Matches are made randomly; the probability that any given agent is matched in any given period with an agent in a given set is proportional to the measure of that set in the market in that period. Notice that since the probability of an agent being matched is less than one, in every period there are agents who have never been matched. Thus even though agents leave the market as time passes, at any finite time a positive measure of every type remains.

Bargaining Once a match is established, each party learns the type (i.e. utility function and initial bundle) and current bundle of his opponent. The members of the match then conduct a short bargaining session. First, one of them is selected to propose a vector z of goods, to be transferred to him from his opponent. (That is, an agent who holds the bundle x and proposes the trade z will hold the bundle $x+z$ if his proposal is accepted.) This vector will typically contain positive and negative elements; it must have the property that it is feasible, in the sense that the bundles held by both parties after the exchange are nonnegative. The probability of each party being selected to make a

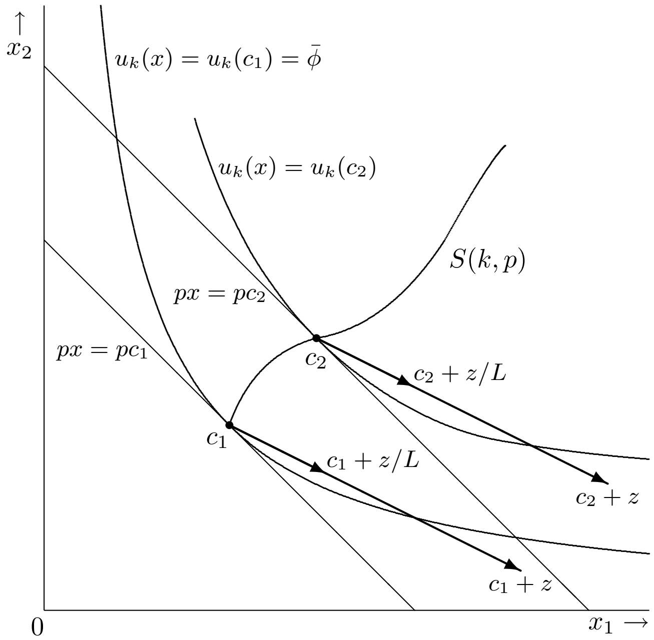


Figure 8.1 An illustration of Assumption 2 on the utility functions.

proposal is 1/2, independent of all past events. After a proposal is made, the other party either accepts or rejects the offer.

Exit In the event an agent rejects an offer, he chooses whether or not to stay in the market. An agent who makes an offer, accepts an offer, or who is unmatched, must stay in the market until the next period: he may not exit. An agent who exits obtains the utility of the bundle he holds at that time.

8.5 Market Equilibrium

A *strategy* for an agent is a plan that prescribes his bargaining behavior for each period, each bundle he currently holds, and each type and current bundle of his opponent. An agent's bargaining behavior is specified by the offer to be made in case he is chosen to be the proposer and, for each possible offer, one of the actions "accept", "reject and stay", or "reject and exit".

An assumption that leads to this definition of a strategy is that each agent observes the index of the period, his current bundle, and the current bundle and type of his opponent, but no past events. Events in the life of the agent (like the type of agents he met in the past, the offers that were made, and the sequence of trades) cannot affect his behavior except insofar as they influence his current bundle. Gale (1986a, Proposition 1) derives the restriction from more primitive assumptions. The idea is the following. Given that there is a continuum of agents, the probability of an agent meeting any particular individual is zero, so that an agent can learn from his personal history about only a finite number of other agents—a set of measure zero. Further, the matching technology forces partners to separate at the end of each period. Thus even if an agent records the entire list of past events, there is no advantage in conditioning his strategy on this information.

We restrict attention to the case in which all agents of a given type use the same strategy. As trade occurs, the bundle held by each agent changes. Different agents of the same type, even though they use the same strategy, may execute different trades. Thus the number of different bundles held by agents may increase. However, the number of different bundles held by agents is finite at all times. Thus in any period the market is given by a finite list $(k_i, c_i, \nu_i)_{i=1,\dots,I}$, where ν_i is the measure of agents who are still in the market, currently hold the bundle c_i, and are of type k_i. We call such a list a *state of the market.* We say that an agent of type k who holds the bundle c is *characterized by* (k, c).

With each K-tuple σ of strategies is associated a state of the market $\rho(\sigma, t)$ in each period t. Although each agent faces uncertainty, the presence of a continuum of agents allows us to define ρ in a deterministic fashion. For example, since in each period the probability that any given agent is matched is α, we take the fraction of agents with any given characteristic who are matched to be precisely α.

Formally, $\rho(\sigma, t+1)$ is generated from $\rho(\sigma, t) = (k_i, c_i, \nu_i)_{i=1,\dots,I}$ by the following transition rules. The set of agents characterized by (k_j, c_j) who are matched with agents characterized by (k_h, c_h) and are selected to make an offer has measure $\alpha\nu_j\nu_h/2\sum_{i=1}^{I}\nu_i$. If σ instructs these agents to offer a trade z that, according to σ, is accepted, then the measure $\alpha\nu_j\nu_h/2\sum_{i=1}^{I}\nu_i$ of agents is transferred from (k_j, c_j) to $(k_j, c_j + z)$, and the measure $\alpha\nu_j\nu_h/2\sum_{i=1}^{I}\nu_i$ of agents is transferred from (k_h, c_h) to $(k_h, c_h - z)$. If σ instructs the responders to reject z and exit, then the measure of agents characterized by (k_h, c_h) is reduced by $\alpha\nu_j\nu_h/2\sum_{i=1}^{I}\nu_i$. Otherwise the measures of agents remain the same.

As an illustration of the determination of $\rho(\sigma, t)$, consider a market in which there are two types, each comprising half of the population. Both

types have the same utility function. There are two goods; each agent of type 1 initially owns the bundle $(2,0)$, while each agent of type 2 owns the bundle $(0,2)$. Suppose that the agents use the following pair of strategies. An agent of type 1 offers and accepts only the trade $(-1,1)$ whenever he holds the bundle $(2,0)$; in all other cases he offers $(0,0)$ and rejects all offers. An agent of type 2 offers and accepts only the trade $(1,-1)$ whenever he holds the bundle $(0,2)$; in all other cases he offers $(0,0)$ and rejects all offers. An agent leaves the market if and only if he holds the bundle $(1,1)$, is matched with a partner, and is chosen to respond to an offer.

In any period, the bundle held by each agent is $(2,0)$, $(0,2)$, or $(1,1)$. Suppose that in period t the measures of agents holding these three bundles are p, q, and r. Let $s = p + q + r$. The measures of agents holding these bundles in period $t+1$ can be found as follows. The measure αr of those holding $(1,1)$ will be matched in period $t+1$; the measure $\alpha r/2$ will be chosen to respond, and hence will leave the market. The remainder of those holding $(1,1)$ (the measure $r(1-\alpha)/2$) will stay in the market through period $t+1$, making the null offer $(0,0)$ if matched. Of the agents holding $(2,0)$, the measure $\alpha pq/s$ will be matched with agents holding $(0,2)$, and will trade and join the set of agents holding $(1,1)$. The remainder will retain $(2,0)$. Thus the total measure of agents holding $(2,0)$ in period $t+1$ is $p(1-\alpha q/s)$. Similarly, the total measure of agents holding $(0,2)$ in period $t+1$ is $q(1-\alpha p/s)$. The total measure of agents holding $(1,1)$ in period $t+1$ is $2\alpha pq/s + r(1-\alpha/2)$.

We emphasize that although we take the evolution of the state of the market to be deterministic, each agent still faces a nondegenerate stochastic process. Given a strategy profile σ, for all pairs (k,c) the state of the market $\rho(\sigma,t)$ induces a well-defined probability that any agent will be matched in period t with an agent characterized by (k,c).

The notion of equilibrium we use is the following.

Definition 8.3 A *market equilibrium* is a K-tuple σ^* of strategies, one for each type, each of which satisfies the following condition for any trade z, bundles c and c', type k, and period t. The behavior prescribed by each agent's strategy from period t on is optimal, given that in period t the agent holds c and has either to make an offer or to respond to the offer z made by his opponent, who is of type k and holds the bundle c', given the strategies of the other types, and given that the agent believes that the state of the market is $\rho(\sigma^*,t)$.

This notion of equilibrium is not directly equivalent to any game-theoretic notion. However, as in the previous model, it is closely related to the notion of sequential equilibrium. Each agent's strategy has to be optimal

after *every* event, including events that are inconsistent with the equilibrium. (These events are: (1) being matched in period t with an agent of type k holding a bundle c when no agent of type k holds c in period t if all agents follows σ^*; (2) being confronted with an offer that the opponent does not make if he adheres to σ^*; (3) having an offer rejected when σ^* calls for the opponent to accept; (4) making a move that is different from that dictated by σ^*.) In order to test the optimality of his strategy, an agent must form a belief about the state of the market, which determines the probabilities with which he meets the various types of agents. If no unexpected event has occurred up to period t, then the equilibrium state of the market in period t, namely $\rho(\sigma^*, t)$, provides this belief. However, once an event that is inconsistent with equilibrium has occurred, an agent must make a conjecture about the current state of the market. The definition of equilibrium requires that each agent believe that, after any sequence of events, the state of the market is the same as it is in equilibrium. This excludes the possibility that an agent interprets out-of-equilibrium behavior by other agents as a signal that the behavior of a *positive* measure of agents was different than in equilibrium, so that the state of the market has changed. This assumption is close to that of the previous model.

8.6 Characterization of Market Equilibrium

An *allocation* is a K-tuple of bundles $(x_1, \ldots, x_K)$ for which $\sum_{k=1}^{K} n_k x_k = \sum_{k=1}^{K} n_k \omega_k$. An allocation $(x_1, \ldots, x_K)$ is *competitive* if there exists a price vector $p \in \mathcal{R}^m_{++}$ such that for all k the bundle x_k maximizes u_k over the budget set $\{x \in X_k : px \le p\omega_k\}$.

The result below establishes a close relationship between competitive allocations and the allocations induced by market equilibria. Before stating this result we need to introduce some terminology. Suppose that the market equilibrium calls for agents characterized by (k, c) who are matched in period t with agents characterized by (k', c') to reject some offer z and leave the market. Then we say that all agents characterized by (k, c) are *ready to leave the market* in period t.

Proposition 8.4 *For every market equilibrium there is a competitive allocation $(x_1, \ldots, x_K)$ such that each agent of type k $(= 1, \ldots, K)$ leaves the market with the bundle x_k with probability one.*

Proof. Consider a market equilibrium; all of our statements are relative to this equilibrium. All agents of type k who hold the bundle c at the beginning of period t (before their match has been determined) face the same probability distribution of future trading opportunities. Thus in the

equilibrium all such agents have the same expected utility; we denote this utility by $V_k(c,t)$.

Step 1. $V_k(c,t) \geq u_k(c)$ for all values of k, c, and t.

Proof. Suppose that an agent of type k who holds the bundle c in period t makes the null offer whenever he is matched and is chosen to propose a trade, and rejects every offer and leaves the market when he is matched and chosen to respond. Since he is matched and chosen to respond to an offer in finite time with probability one, this strategy guarantees him a payoff of $u_k(c)$. (Recall that all agents are indifferent about the timing of consumption.) Thus $V_k(c,t) \geq u_k(c)$.

Step 2. $V_k(c,t) \geq V_k(c,t+1)$ for all values of k, c, and t.

Proof. The assertion follows from the fact that by proposing the null trade and rejecting any offer and staying in the market, any agent in the market in period t is sure of staying in the market until period $t+1$ with his current bundle.

Step 3. For an agent of type k who holds the bundle c and is ready to leave the market in period t we have $V_k(c,t+1) = u_k(c)$.

Proof. By Step 1 we have $V_k(c,t+1) \geq u_k(c)$. If $V_k(c,t+1) > u_k(c)$ and the circumstances that make the agent leave the market are realized (in which case he would leave with the bundle c), then he is better off by deviating and staying in the market until period $t+1$.

Step 4. Suppose that an agent of type k holds the bundle c and is ready to leave the market in period t. Then it is optimal for him to accept any offer z (of a transfer *from* him *to* the proposer) for which $u_k(c-z) > u_k(c)$.

Proof. If he accepts the offer, then his expected utility $V_k(c-z,t+1)$ in the continuation is at least $u_k(c-z)$ (by Step 1), and this exceeds his expected utility in the continuation if he rejects the offer, which is $V_k(c,t+1) = u_k(c)$ (see Step 3).

Step 5. For any period t and any given agent, the probability that in some future period the agent will be chosen to make an offer to an agent who is ready to leave the market is one.

Proof. Let Q_s be the measure of the set of agents in the market in period s, and let E_s be the measure of the set of agents who are ready to leave the market in period s. The probability that an agent in the market is matched with an agent who is ready to leave is $\alpha E_s/Q_s$, in which case the agent will be chosen with probability $1/2$ to make an offer. Thus the probability of not being able to make an offer to an agent who is ready to

leave is $1-\alpha E_s/2Q_s$. The measure of agents who actually leave is at most $\alpha E_s/2$. (Recall that an agent who is ready to leave does so only under some circumstances, not necessarily whenever he has to respond to an offer.) Hence $Q_s - \alpha E_s/2 \leq Q_{s+1}$, so that $1-\alpha E_s/2Q_s \leq Q_{s+1}/Q_s$. Thus the probability of not being able to make an offer from period t through period s to an agent who is ready to leave the market, where $s > t$, is at most Q_{s+1}/Q_t. Since the utility of staying in the market forever is $-\infty$, $Q_s \to 0$ as $s \to \infty$, and thus the probability that an agent will be able to make an offer in some future period to an agent who is ready to leave the market is one.

Step 6. There is a vector $p \in \mathcal{R}^m_{++}$, unique up to multiplication by a nonnegative scaler, such that, for all k, if each member of a set of agents of positive measure of type k leaves the market in some period with the bundle c, then the tangent to the indifference curve $\{x \in X_k \colon u_k(x) = u_k(c)\}$ at c is $\{x \colon px = pc\}$ (i.e. $pz > 0$ for all z such that $u_k(c+z) > u_k(c)$).

Proof. Suppose that each member of a set of positive measure of agents of type k_1 leaves the market in period t_1 with the bundle c_1, and each member of a set of positive measure of agents of type k_2 leaves the market in period t_2 with the bundle c_2. Assume, contrary to the claim, that the tangent to the indifference curve $\{x \in X_{k_1} \colon u_{k_1}(x) = u_{k_1}(c_1)\}$ at c_1 is different from the tangent to the indifference curve $\{x \in X_{k_2} \colon u_{k_2}(x) = u_{k_2}(c_2)\}$ at c_2. Then (by the assumption that each indifference curve has a unique tangent at every point) there is a trade z between an agent of type k_1 holding the bundle c_1 and an agent of type k_2 holding the bundle c_2 that makes both agents better off. More precisely, $c_1 + z \in X_{k_1}$, $c_2 - z \in X_{k_2}$, $u_{k_1}(c_1 + z) > u_{k_1}(c_1)$, and $u_{k_2}(c_2 - z) > u_{k_2}(c_2)$.

First assume that $t_1 < t_2$. Consider an agent of type k_1 who holds the bundle c_1 in period t_1. By our hypothesis he is ready to leave the market. We will show that the following is a profitable deviation. Instead of leaving the market, he stays until period t_2 (by proposing the null trade and rejecting all offers as necessary). In period t_2 there is a positive probability that he is matched with an agent of type k_2 who holds c_2 (and thus is ready to leave the market). In this event he proposes the mutually beneficial trade z. In every other event he departs from the market at the first opportunity. By Step 4 the agent of type k_2 accepts the offer, so that the agent of type k_1 either achieves the bundle $c_1 + z$ in period t_2 (with positive probability) or holds the bundle c_1 in that period. Thus the agent of type k_1 achieves an expected utility in excess of $u_{k_1}(c_1)$, so that the deviation is profitable.

If $t_1 = t_2 = t$ then an agent of type k_1 who faces the circumstances in which he plans to leave the market can deviate from the equilibrium and

postpone his departure by one period. In period $t + 1$ there is a positive probability that he will be matched with an agent of type k_2 who was ready to leave the market in period t but did not do so. Suppose the agent of type k_1 offers the trade z to such an agent of type k_2. If the latter accepts this offer then by Step 1 that agent's expected utility in the continuation is at least $u_{k_2}(c_2 - z)$, while if he rejects the offer then he either leaves the market with the bundle c_2 or enters period $t + 2$ with that bundle. But $V_{k_2}(c_2, t + 2) \leq V_{k_2}(c_2, t + 1) = u_{k_2}(c_2)$ by Steps 2 and 3. So the fact that $u_{k_2}(c_2 - z) > u_{k_2}(c_2)$, and the requirement that each agent's strategy prescribe optimal actions after every history, demand that the agent of type k_2 accept the offer. Hence, as in the previous case, the agent of type k_1 has a profitable deviation from his purported equilibrium strategy.

We conclude that the tangents to the indifference curves of agents who leave the market, at the bundles with which they depart, coincide.

Step 7. Let p be the vector defined in Step 6. Then for all k, c, and t we have $V_k(c, t) \geq \max_{x \in X_k}\{u_k(x) \colon px \leq pc\}$.

Proof. Assume to the contrary that $V_\kappa(c, t) < \max_{x \in X_\kappa}\{u_\kappa(x) \colon px \leq pc\}$ for some κ, c, and t. Then there is a vector z such that $V_\kappa(c, t) < u_\kappa(c + z)$ and $pz < 0$. (See Figure 8.2.) We shall argue that an agent of type κ who holds the bundle c has a deviation that yields him the utility $u_\kappa(c + z)$. By Assumption 2 (p. 158) for each $k = 1, \ldots, K$ there exists a positive integer L_k such that $u_k(c_k - z/L_k) > u_k(c_k)$ whenever c_k is a bundle with which agents of type k leave the market (using Step 6). By Step 4, an agent of type k who is ready to leave the market thus accepts an offer of the trade $-z/L_k$. Hence there exists a positive integer L such that all agents (of whatever type) who are ready to leave the market would accept the trade $-z/L$ before doing so. Now, by Step 5 the probability that in some future period a given agent of type κ will be able to make an offer to an agent who is ready to leave the market is one. Thus the probability that he will be able to make L such offers is also one. Hence an agent of type κ who holds the bundle c in period t can profitably deviate from his original strategy and with certainty carry out L trades of z/L before he leaves the market, thereby attaining the utility $u_\kappa(c + z)$, which exceeds $V_\kappa(c, t)$.

Step 8. For an agent of type k who leaves the market with the bundle c we have $u_k(c) = \max_{x \in X_k}\{u_k(x) \colon px \leq p\omega_k\}$.

Proof. By Step 7 we have $V_k(\omega_k, 0) \geq \max_{x \in X_k}\{u_k(x) \colon px \leq p\omega_k\}$. Let $\mathbf{y}_k$ be the random bundle with which an agent of type k leaves the market. We show that for all k the random variable $\mathbf{y}_k$ is degenerate and $V_k(\omega_k, 0) = \max_{x \in X_k}\{u_k(x) \colon px \leq p\omega_k\}$. Assume this is not so for

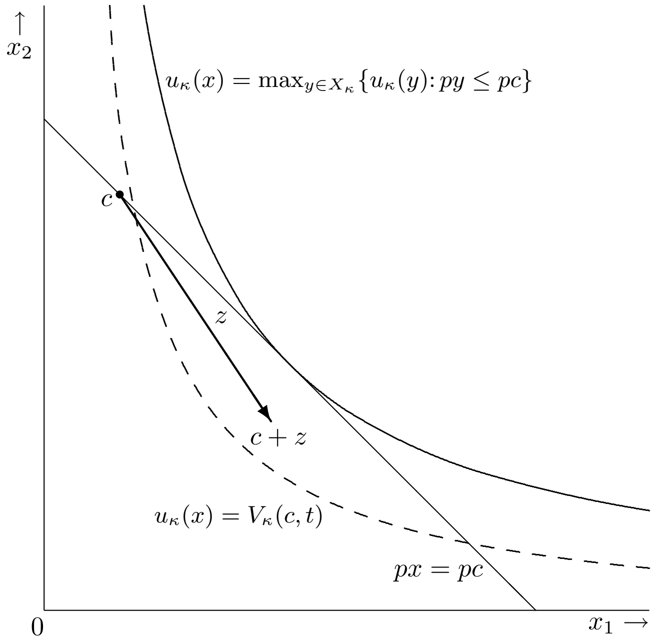


Figure 8.2 A vector z for which $V_\kappa(c, t) < u_\kappa(c + z)$ and $pz < 0$.

$k = \kappa$. By the strict concavity of u_k and Jensen's inequality we have $V_k(\omega_k, 0) = \mathrm{E}[u_k(\mathbf{y}_k)] \le u_k(\mathrm{E}[\mathbf{y}_k])$ (where E is the expectation operator), with strict inequality unless $\mathbf{y}_k$ is degenerate. Let $y_k = \mathrm{E}[\mathbf{y}_k]$. Hence $u_k(y_k) \ge \max_{x \in X_k}\{u_k(x): px \le p\omega_k\}$, with strict inequality for $k = \kappa$. Therefore $py_k \ge p\omega_k$ for all k, and $py_\kappa > p\omega_\kappa$. Thus $p\sum_{k=1}^K n_k y_k > p\sum_{k=1}^K n_k \omega_k$, contradicting the condition $\sum_{k=1}^K n_k y_k = \sum_{k=1}^K n_k \omega_k$ for $(y_1, \ldots, y_K)$ to be an allocation. □

Note that Assumption 2 (p. 158) is used in Step 7. It is used to show that if $pz < 0$ then there is a trade in the direction $-z$ that makes any agent who is ready to leave the market better off. Thus, by executing a sequence of such trades, an agent who holds the bundle c is assured of eventually obtaining the bundle $c - z$. Suppose the agents' preferences do not satisfy Assumption 2. Then the curvature of the agents' indifference curves at the bundles with which they exit from the market in period t might increase with t, in such a way that the exiting agents are willing to accept only a

sequence of successively smaller trades in the direction $-z$, a sequence that never adds up to z itself.

Two arguments are central to the proof. First, the allocation associated with the bundles with which agents exit is efficient (Step 6). The idea is that if there remain feasible trades between the members of two sets of agents that make the members of both sets better off, then by waiting sufficiently long each member of one set is sure of meeting a member of the other set, in which case a mutually beneficial trade can take place. Three assumptions are important here. First, no agent is impatient. Every agent is willing to wait as long as necessary to execute a trade. Second, the matching technology has the property that if in some period there is a positive measure of agents of type k holding the bundle c, then in *every* future period there will be a positive measure of such agents, so that the probability that any other given agent meets such an agent is positive. Third, an agent may not leave the market until he has rejected an offer. This gives every agent a chance to make an offer to an agent who is ready to leave the market. If we assume that an agent can leave the market whenever he wishes then we cannot avoid inefficient equilibria in which all agents leave the market simultaneously, leaving gains from trade unexploited.

The second argument central to the proof is contained in Step 7. Consider a market containing two types of agents and two goods. Suppose that the bundles with which the members of the two types exit from the market leave no opportunities for mutually beneficial trade unexploited. Given the matching technology, in every period there will remain agents of each type who have never been matched and hence who still hold their initial bundles. At the same time, after a number of periods some agents will hold their final bundles, ready to leave the market. If the final bundles are not competitive, then for one of the types—say type 1—the straight line joining the initial bundle and the final bundle intersects the indifference curve through the final bundle. This means that there is some trade z with the property that $u_1(\omega_1 + Lz) > u_1(x_1)$ for some integer L, where x_1 is the final bundle of an agent of type 1, and $u_1(x_1 - z) > u_1(x_1)$. Put differently, a number of executions of z makes an agent of type 1 currently holding the initial bundle better off than he is when he holds the final bundle, and a single execution of $-z$ makes an agent of type 1 who is ready to leave the market better off. Given the matching technology, any agent can (eventually) meet as many agents of type 1 who are ready to leave as he wishes. Thus, given that the matching technology forces some agents to achieve their final bundles before others (rather than all of them achieving the final bundles simultaneously), there emerge unexploited opportunities for trade whenever the final outcome is not competitive, even when it is ef-

ficient. Once again we see the role of the three assumptions that the agents are patient, the matching technology leaves a positive measure unmatched in every period, and an agent cannot exit until he has rejected an offer. Another assumption that is significant here is that each agent can make a sequence of transactions before leaving the market. This assumption increases the forces of competition in the market, since it allows an agent to exploit the opportunity of a small gain from trade without prejudicing his chances of participating in further transactions.

8.7 Existence of a Market Equilibrium

Proposition 8.4 leaves open the question of the existence of a market equilibrium. Gale (1986b) studies this issue in detail and establishes a converse of Proposition 8.4: to every competitive equilibrium there is a corresponding market equilibrium. (Thus, in particular, a market equilibrium exists.) We do not provide a detailed argument here. Rather we consider two cases in which a straightforward argument can be made.

First consider a modification of the model in which agents may make "short sales"—that is, agents may hold negative amounts of goods, so that any trade is feasible. This case avoids some difficulties associated with the requirement that trades be feasible and illustrates the main ideas. (It is studied by McLennan and Sonnenschein (1991).) Assume that for every bundle c, type k, and price vector p, the maximizer of $u_k(x)$ over $\{x : px \leq pc\}$ is unique, and let $\hat{z}(p,c,k)$ be the difference between this maximizer and c; we refer to $\hat{z}(p,c,k)$ as the *excess demand* at the price vector p of an agent characterized by (k,c). If $\hat{z}(p,c,k) = 0$ then an agent characterized by (k,c) holds the bundle (c) that maximizes his utility at the price vector p. Let p^* be the price vector corresponding to a competitive equilibrium of the market. Consider the strategy profile in which the strategy of an agent characterized by (k,c) is the following. Propose the trade $\hat{z}(p^*,k,c)$. If $\hat{z}(p^*,k,c) \neq 0$ then accept an offer[1] z if $p^*(-z) \geq 0$; otherwise reject z and stay in the market. If $\hat{z}(p^*,k,c) = 0$ then accept an offer z if $p^*(-z) > 0$; otherwise reject z and leave the market. The outcome of this strategy profile is that each agent eventually leaves the market with his competitive bundle (the bundle that maximizes his utility over his budget set at the price p^*). If all other agents adhere to the strategy profile, then any given agent accepts any offer he is faced with; his proposal to trade his excess demand is accepted the first time he is matched and chosen to be the proposer, and he leaves the market in the next period in which he is matched and chosen to be the responder.

[1] That is, a trade after which the agent holds the bundle $c - z$.

We claim that the strategy profile is a market equilibrium. It is optimal for an agent to accept any trade that results in a bundle that is worth not less than his current bundle, since with probability one he will be matched and chosen to propose in the future, and in this event his proposal to trade his excess demand will be accepted. It is optimal for an agent to reject any trade that results in a bundle that is worth less than his current bundle, since no agent accepts any trade that decreases the value of his bundle. Finally, it is optimal for an agent to propose his excess demand, since this results in the bundle that gives the highest utility among all the trades that are accepted.

We now return to the model in which in each period each agent must hold a nonnegative amount of each good. In this case the trading strategies must be modified to take into account the feasibility constraints. We consider only the case in which there are two goods, the market contains only two types of equal measure, and the initial allocation is not competitive. Then for any competitive price p^* we have $\hat{z}(p^*, 1, \omega_1) = -\hat{z}(p^*, 2, \omega_2) \neq 0$. Consider the strategy profile in which the strategy of an agent characterized by (k, c) is the following.

Proposals Propose the maximal trade in the direction of the agent's optimal bundle that does not increase or change the sign of the responder's excess demand. Precisely, if matched with an agent characterized by (k', c') and if $\hat{z}_1(p^*, k, c)$ has the same sign as $\hat{z}_1(p^*, k', c')$ (where the subscript indicates good 1), then propose $z = 0$. Otherwise, propose the trade $\hat{z}(p^*, k, c)$ if $|\hat{z}(p^*, k, c)| \leq |\hat{z}(p^*, k', c')|$, and the trade $-\hat{z}(p^*, k', c')$ if $|\hat{z}(p^*, k, c)| > |\hat{z}(p^*, k', c')|$, where $|x|$ is the Euclidian norm of x.

Responses If $\hat{z}(p^*, k, c) \neq 0$ then accept an offer z if $p^*(-z) > 0$, or if $p^*(-z) = 0$ and $\hat{z}_i(p^*, k, c - z)$ has the same sign as, and is smaller than $\hat{z}_i(p^*, k, c)$ for $i = 1, 2$. Otherwise reject z and stay in the market. If $\hat{z}(p^*, k, c) = 0$ then accept an offer z if $p^*(-z) > 0$; otherwise reject z and leave the market.

As in the previous case, the outcome of this strategy profile is that each agent eventually leaves the market with the bundle that maximizes his utility over his budget set at the price p^*. If all other agents adhere to the strategy profile, then any given agent realizes his competitive bundle the first time he is matched with an agent of the other type; until then he makes no trade. The argument that the strategy profile is a market equilibrium is very similar to the argument for the model in which the feasibility constraints are ignored. An agent characterized by (k, c) is assured of eventually achieving the bundle that maximizes u_k over $\{x \in X_k : px \leq pc\}$,

since he does so after meeting only a finite number of agents of one of the types who have never traded (since any such agent has a nonzero excess demand), and the probability of such an event is one.

8.8 Market Equilibrium and Competitive Equilibrium

Propositions 8.2 and 8.4 show that the noncooperative models of decentralized trade we have defined lead to competitive outcomes. The first proposition, and the arguments of Gale (1986b), show that the converse of the results are also true: every distribution of the goods that is generated by a competitive equilibrium can be attained as the outcome of a market equilibrium.

In both models the technology of trade and the agents' lack of impatience give rein to competitive forces. If, in the first model, a price below 1 prevails, then a seller can push the price up by waiting (patiently) until he has the opportunity to offer a slightly higher price; such a price is accepted by a buyer since otherwise he will be unable, with positive probability, to purchase the good. If, in the second model, the allocation is not competitive, then an agent is able to wait (patiently) until he is matched with an agent to whom he can offer a mutually beneficial trade.

An assumption that is significant in the two models is that agents cannot develop personal relationships. They are anonymous, are forced to separate at the end of each bargaining session, and, once separated, are not matched again. In Chapter 10 we will see that if the agents have personal identities then the competitive outcome does not necessarily emerge.

Notes

The model of Section 8.2 is closely related to the models of Binmore and Herrero (1988a) and Gale (1987, Section 5), although the exact form of Proposition 8.2 appears in Rubinstein and Wolinsky (1990). The model of Section 8.4 and the subsequent analysis is based on Gale (1986c), which is a simplification of the earlier paper Gale (1986a). The existence of a market equilibrium in this model is established in Gale (1986b).

Proposition 8.2 is related to Gale (1987, Theorem 1), though Gale deals with the limit of the equilibrium prices when $\delta \to 1$, rather than with the limit case $\delta = 1$ itself. Gale's model differs from the one here in that there is a finite number of types of agents (distinguished by different reservation prices), and a continuum of agents of each type. Further, each agent can condition his behavior on his entire personal history. However, given the matching technology and the fact that each pair must separate at the end of each period, the only information relevant to each agent is the time

and the names of the agents remaining in the market, as we assumed in Proposition 8.2. Thus we view Proposition 8.2 as the analog of Gale's theorem in the case that the market contains a finite number of agents.

Binmore and Herrero (1988a) investigate alternative information structures and define a solution concept that leads to the same conclusion about the relation between the sets of market equilibria and competitive equilibria as the models we have described. The relation between Proposition 8.4 and the theory of General Equilibrium is investigated by McLennan and Sonnenschein (1991), who also prove a variant of the result under the assumption that the behavior dictated by the strategies does not depend on time. Gale (1986e) studies a model in which the agents—workers and firms—are asymmetrically informed. Workers differ in their productivities and in their payoffs outside the market under consideration. These productivities and payoffs are not known by the firms and are positively correlated, so that a decrease in the offered wage reduces the quality of the supply of workers. Gale examines the nature of wage schedule offered in equilibrium.

CHAPTER 9

The Role of the Trading Procedure

9.1 Introduction

In this chapter we focus on the role of the trading procedure in determining the outcome of trade. The models of markets in the previous three chapters have in common the following three features.

1. The bargaining is always bilateral. All negotiations take place between two agents. In particular, an agent is not allowed to make offers simultaneously to more than one other agent.

2. The termination of an unsuccessful match is exogenous. No agent has the option of deciding to stop the negotiations.

3. An agreement is restricted to be a price at which the good is exchanged. Other agreements are not allowed: a pair of agents cannot agree that one of them will pay the other to leave the market, or that they will execute a trade only under certain conditions.

The strategic approach has the advantage that it allows us to construct models in which we can explore the role of these three features.

As in other parts of the book, we aim to exhibit only the main ideas in the field. To do so we study several models, in all of which we make the following assumptions.

Goods A single indivisible good is traded for some quantity of a divisible good ("money").

Time Time is discrete and is indexed by the nonnegative integers.

Economic Agents In period 0 a single seller, whom we refer to as S, and two buyers, whom we refer to as B_H and B_L, enter the market. The seller owns one unit of the indivisible good. The two buyers have reservation values for the good of v_H and v_L, respectively, where $v_H \geq v_L > 0$. No more agents enter the market at any later date (cf. Model B in Chapter 6). All three agents have time preferences with a constant discount factor of $0 < \delta < 1$. An agreement on the price p in period t yields a payoff of $\delta^t p$ for the seller and of $\delta^t(v - p)$ for a buyer with reservation value v. If an agent does not trade then his payoff is zero. When uncertainty is involved we assume that the agents maximize their expected utilities.

Information All agents have full information about the history of the market at all times: the seller always knows the buyer with whom she is matched, and every agent learns about, and remembers, all events that occur in the market, including the events in matches in which he does not take part.

In a market containing only S and B_H, the price at which the good is sold in the unique subgame perfect equilibrium of the bargaining game of alternating offers in which S makes the first offer is $v_H/(1+\delta)$. We denote this price by p_H^*.

When bargaining with B_H, the seller can threaten to trade with B_L, so that it appears that the presence of B_L enhances her bargaining position. However, the threat to trade with B_L may not be credible, since the surplus available to S and B_L is lower than that available to S and B_H. Thus the extent to which the seller can profit from the existence of B_L is not clear; it depends on the exact trading procedure.

We start, in Section 9.2, with a model in which the three features mentioned at the beginning of this section are retained. As in the previous three chapters we assume that the matching process is random and is given exogenously. A buyer who rejects an offer runs the risk of losing the seller and having to wait to be matched anew. We show that if $v_H = v_L$ then this fact improves the seller's bargaining position: the price at which the good is sold exceeds p_H^*.

Next, in Section 9.3, we study a model in which the seller can make an offer that is heard simultaneously by the two buyers. We find that if v_H is not too large and δ is close to 1, then once again the presence of B_L increases the equilibrium price above p^*_H.

In Section 9.4 we assume that in each period the seller can choose the buyer with whom to negotiate. The results in this case depend on the times at which the seller can switch to a new buyer. If she can switch only after she rejects an offer, then the equilibrium price is precisely p^*_H: in this case a threat by S to abandon B_H is not credible. If the seller can switch only after the buyer rejects an offer, then there are many subgame perfect equilibria. In some of these, the equilibrium price exceeds p^*_H.

Finally, in Section 9.5 we allow B_H to make a payment to B_L in exchange for which B_L leaves the market, and we allow the seller to make a payment to B_L in exchange for which B_L is committed to buying the good at the price v_L in the event that S does not reach agreement with B_H. The equilibrium payoffs in this model coincide with those predicted by the Shapley value; the equilibrium payoff of the seller exceeds p^*_H.

We see that the results we obtain are sensitive to the precise characteristics of the trading procedure. One general conclusion is that only when the procedure allows the seller to effectively commit to trade with B_L in the event she does not reach agreement with B_H does she obtain a price that exceeds p^*_H.

9.2 Random Matching

At the beginning of each period the seller is randomly matched with one of the two buyers, and one of the matched parties is selected randomly to make a proposal. Each random event occurs with probability 1/2, independent of all past events. The other party can either accept or reject the proposal. In the event of acceptance, the parties trade, and the game ends. In the event of rejection, the match dissolves, and the seller is (randomly) matched anew in the next period. Note that the game between the seller and the buyer with whom she is matched is similar to the model of alternating offers with breakdown that we studied in Section 4.2 (with a probability of breakdown of 1/2). The main difference is that the payoffs of the agents in the event of breakdown are determined endogenously rather than being fixed.

9.2.1 The Case $v_H = v_L$

Without loss of generality we let $v_H = v_L = 1$. The game has a unique subgame perfect equilibrium, in which the good is sold to the first buyer to be matched at a price close to the competitive price of 1.

Proposition 9.1 *If $v_H = v_L = 1$ then the game has a unique subgame perfect equilibrium, in which the good is sold immediately at the price $p_s = (2-\delta)^2/(4-3\delta)$ if the seller is selected to make the first offer, and at the price $p_b = \delta(2-\delta)/(4-3\delta)$ if the matched buyer is selected to make the first offer. These prices converge to 1 as δ converges to 1.*

Proof. Define M_s and m_s to be the supremum and the infimum of the seller's payoff over all subgame perfect equilibria of the game. Similarly, define M_b and m_b to be the corresponding values for either of the buyers *in the same game*. Four equally probable events may occur at the beginning of each period. Denoting by i/j the event that i is selected to make an offer to j, these events are S/B_H, B_H/S, S/B_L, and B_L/S.

Step 1. $M_s \geq (2(1-\delta m_b) + 2\delta M_s)/4$ and $m_b \leq (1-\delta M_s + \delta m_b)/4$.

Proof. For every subgame perfect equilibrium that gives j a payoff of v we can construct a subgame perfect equilibrium for the subgame starting with the event i/j such that agreement is reached immediately, j's payoff is δv and i's payoff is $1-\delta v$. The inequalities follow from the fact that there exists a subgame perfect equilibrium such that after each of the events S/B_I the good is sold at a price arbitrarily close to $1-\delta m_b$, and after each of the events B_I/S the good is sold at a price arbitrarily close to δM_s.

Step 2. $m_b = (1-\delta)/(4-3\delta)$ and $M_s = (2-\delta)/(4-3\delta)$.

Proof. The seller obtains no more than δM_s when she has to respond, and no more than $1-\delta m_b$ when she is the proposer. Hence $M_s \leq (2\delta M_s + 2(1-\delta m_b))/4$. Combined with Step 1 we obtain $M_s = (2\delta M_s + 2(1-\delta m_b))/4$. Similarly, a buyer obtains at least $1-\delta M_s$ when he is matched and is chosen to be the proposer, and at least δm_b when he is matched and is chosen to respond. Therefore $m_b \geq (1-\delta M_s + \delta m_b)/4$, which, combined with Step 1, means that $m_b = (1-\delta M_s + \delta m_b)/4$. The two equalities imply the result.

Step 3. $M_b \leq 1 - m_b - m_s$.

Proof. This follows from the fact that the most that a buyer gets in equilibrium does not exceed the surplus minus the sum of the minima of the two other agents' payoffs.

Step 4. $M_s = m_s = (2-\delta)/(4-3\delta)$ and $M_b = m_b = (1-\delta)/(4-3\delta)$.

Proof. If the seller is the responder then she obtains at least δm_s, and if she is the proposer then she obtains at least $1-\delta M_{b.By Step\ 3 we have} 1-\delta M_b \geq 1-\delta(1-m_b-m_s)$, so that $m_s \geq [2\delta m_s + 2(1-\delta(1-m_b-m_s))]/4$, which implies that $m_s \geq 1/2 + \delta m_b/[2(1-\delta)] = 1/2 + \delta/[2(4-3\delta)] = M_s$. Finally, we have $M_b \leq 1 - m_b - m_s = (1-\delta)/(4-3\delta) = m_b$.

By the same argument as in the proof of Theorem 3.4 it follows that there is a unique subgame perfect equilibrium in which the seller always proposes the price $1 - \delta M_b = p_s$, and each buyer always offers the price $\delta M_s = p_b$. □

Note that the technique used in the proof of Step 1 is different from that used in the proofs of Steps 1 and 2 of Theorem 3.4. Given a collection of subgame perfect equilibria in the subgames starting in the second period we construct a subgame perfect equilibrium for the game starting in the first period. This line of argument is useful in other models that are similar to the one here.

So far we have assumed that a match may be broken after any offer is rejected. If instead a match may be broken only after the seller rejects an offer, then the unique subgame perfect equilibrium coincides with that in the game in which the seller faces a single buyer (and the proposer is chosen randomly at the start of each period). The prices the agents propose thus converge to $1/2$ as δ converges to 1. On the other hand, if a match may be broken only after a buyer rejects an offer, then there is a unique subgame perfect equilibrium, which coincides with the one given in Proposition 9.1. This leads us to a conclusion about how to model competitive forces. If we want to capture the pressure on the price caused by the presence of more than one buyer, we must include in the model the risk that a match may be broken after the buyer rejects an offer; it is not enough that there be this risk only after the seller rejects an offer.

We now consider briefly the case in which the probability that a match terminates after an offer is rejected is one, rather than $1/2$: that is, the case in which the seller is matched in alternate periods with B_H and B_L. Retaining the assumption that the proposer is selected randomly, the game has a unique subgame perfect equilibrium, in which the seller always proposes the price 1, and each buyer always proposes the price $p_b = \delta/(2 - \delta)$. (The equation that determines p_b is $p_b = \delta(1/2 + p_b/2)$.) A buyer accepts the price 1, since if he does not then the good will be sold to the other buyer. When a buyer is selected to make a proposal he is able to extract some surplus from the seller since she is uncertain whether she will be the proposer or the responder in the next match.

If we assume that the matches and the selection of proposer are *both* deterministic, then the subgame perfect equilibrium depends on the order in which the agents are matched and chosen to propose. If the order is S/B_I, B_I/S, S/B_J, B_J/S (for $\{I, J\} = \{L, H\}$), then the unique subgame perfect equilibrium is essentially the same as if there were only one buyer: the seller always proposes the price $1/(1 + \delta)$, while each buyer always proposes $\delta/(1 + \delta)$. If the order is B_I/S, S/B_I, B_J/S, S/B_J then in the unique

subgame perfect equilibrium the seller always proposes the price 1, while each buyer always proposes the price δ. The comparison between these two protocols demonstrates again that in order to model the competition between the two buyers we need to construct a model in which a match is broken after a buyer, rather than a seller, rejects an offer.

9.2.2 *The Case* $v_H > v_L$

We now turn to the case in which the buyers have different reservation values, with $v_H > v_L$. We return to our initial assumptions in this section that each match is terminated with probability 1/2 after a rejection, and that the probability that each of the parties is chosen to be the proposer is also 1/2. If $v_H/2 > v_L$ and δ is close enough to 1, then there is a unique subgame perfect equilibrium in which the good is sold to B_H at a price close to $v_H/2$. The intuition is that the seller prefers to sell the good to B_H at the price that would prevail were B_L absent from the market, so that both the seller and B_H consider the termination of their match to be equally appalling.

We now consider the case $v_H/2 < v_L$. (This is the case we considered in Section 6.5.) In this case, the game does not have a stationary subgame perfect equilibrium if δ is close to 1. The intuition is as follows. Assume that there is a stationary subgame perfect equilibrium in which the seller trades with B_L when she is matched with him, for at least one of the two choices of proposer. The interaction between S and B_H is then the same as in a bilateral bargaining game in which with probability at least 1/4 the match does not continue: negotiations between S and B_H break down, and an agreement is reached between S and B_L. This breakdown is exogenous from the point of view of the interaction between S and B_H. The payoff of B_H of such a breakdown is zero, and some number $u \leq 3v_H/4 + v_L/4 < v_H$ for the seller. The equilibrium price in the bargaining between S and B_H is therefore approximately $(u + v_H)/2$ when δ is close to 1. Since $(u + v_H)/2 > u$, it is thus better for the seller to wait for an opportunity to trade with B_H than to trade with B_L. Thus in no stationary equilibrium does the seller trade with B_L.

Now consider a stationary subgame perfect equilibrium in which the seller trades only with B_H. If δ is close to 1, the surplus v_H is split more or less equally between the seller and B_H. However, given the assumption that $v_L > v_H/2$, buyer B_L should agree to a price between v_L and $v_H/2$, and the seller is better off waiting until she is matched with B_L and has the opportunity to make him such an offer. Therefore there is no stationary equilibrium in which with probability 1 the unit is sold to B_H.

		T_H	T_{HL}
S	proposes to B_H	p^*	p^*
	proposes to B_L	p^*	v_L
	accepts from B_H	$p \geq v_L$	$p \geq v_L$
	accepts from B_L	$p > v_L$	$p \geq v_L$
B_H	proposes	v_L	v_L
	accepts	$p \leq p^*$	$p \leq p^*$
B_L	proposes	v_L	v_L
	accepts	$p \leq v_L$	$p \leq v_L$
	Transitions	Go to T_{HL} if B_H rejects a price $p \leq p^*$.	Go to T_H after any rejection except a rejection of $p \leq p^*$ by B_H.

Table 9.1 A nonstationary subgame perfect equilibrium for the model of Section 9.2.2, under the assumption that $v_L < v_H < 2v_L$. The price p^* is equal to $(4 - 3\delta)v_L/\delta$ $(> v_L)$.

We now describe a nonstationary subgame perfect equilibrium. There are two states, T_H ("trade only with B_H") and T_{HL} ("trade with both B_H and B_L"), and $p^* = (4 - 3\delta)v_L/\delta > v_L$. The initial state is T_H. The strategies are given in Table 9.1.

We now check that this strategy profile is a subgame perfect equilibrium for δ close enough to 1. The price p^* is chosen so that in each state the expected utility of the seller before being matched is v_L/δ. (In state T_H this utility is the number V that satisfies $V = (v_L + p^*)/4 + \delta V/2$; in state T_{HL} it is $p^*/4 + 3v_L/4$.) Therefore in each state the seller is indifferent between selling the good at the price v_L and taking an action that delays agreement. Hence her strategy is optimal.

Now consider the strategy of B_H. It is optimal for him to accept p^* in state T_H since if he rejects it then the state changes to T_{HL}, in which he obtains the good only with probability 1/2. More precisely, if he accepts p^* he obtains $v_H - p^*$, while if he rejects it he obtains $\delta[(1/2) \cdot 0 + (1/4) \cdot (v_H - p^*) + (1/4) \cdot (v_H - v_L)] < v_H - p^*$ if δ is close enough to 1. For a similar reason, B_H cannot benefit by proposing a price less than v_L in either state. It is optimal for him to reject $p > p^*$ in both states since if he accepts it he obtains $v_H - p$, while if he rejects it, the state either remains or becomes T_H, and he obtains close to the average of $v_H - p^*$

and $v_H - v_L$ if δ is close to 1. Precisely, his expected utility before being matched in state T_H is $v_H/(2-\delta) - v_L/\delta$ (the number V that satisfies $V = (1/2)(v_H - (v_L + p^*)/2) + (1/2)\delta V$), which exceeds $v_H - p$ if δ is close enough to 1 and $p > p^*$. Finally, B_L's strategy is optimal since his expected utility is zero in both states.

This equilibrium is efficient, since the good is sold to B_H at the first opportunity. However, the argument shows that there is another subgame perfect equilibrium, in which the initial state is T_{HL} rather than T_H, which is inefficient. In this equilibrium the good is sold to B_L with probability 1/2. We know of no characterization of the set of all subgame perfect equilibria.

9.3 A Model of Public Price Announcements

In this section we relax the assumption that bargaining is bilateral. The seller starts the game by announcing a price, which both buyers hear. Then B_H responds to the offer. If he accepts the offer then he trades with the seller, and the game ends. If he rejects it, then B_L responds to the offer. If both buyers reject the offer, then play passes into the next period, in which *both* buyers simultaneously make counteroffers. The seller may accept one of these, or neither of them. In the latter case, play passes to the next period, in which it is once again the seller's turn to announce a price. Recall that $p_H^* = v_H/(1+\delta)$, the unique subgame perfect equilibrium price in the bargaining game of alternating offers between the seller and B_H in which the seller makes the first offer.

Proposition 9.2 *If $\delta p_H^* < v_L$, then the model of public price announcements has a subgame perfect equilibrium, and in all subgame perfect equilibria the good is sold (to B_H if $v_H > v_L$) at the price $p^* = \delta v_L + (1-\delta)v_H$. If $\delta p_H^* > v_L$ then the game has a unique subgame perfect equilibrium. In this equilibrium the good is sold to B_H at the price p_H^*.*

Thus if the value to the seller of receiving p_H^* with one period of delay is less than v_L then the seller gains from the existence of B_L: $p^* > p_H^*$. The price p^* lies between v_L and v_H; it exceeds v_L if $v_H > v_L$, and converges to v_L as δ converges to 1. By contrast, if the value to the seller of receiving p_H^* with one period of delay exceeds v_L, then the existence of B_L does *not* improve the seller's position. This part of the result is similar to the first part of Proposition 3.5, which shows that the fact that a player has an outside option with a payoff lower than the equilibrium payoff in bilateral bargaining does not affect the bargaining outcome.

Proof of Proposition 9.2. If $\delta p_H^* > v_L$ then there is a subgame perfect equilibrium in which S and B_H behave as they do in the unique subgame

perfect equilibrium of the bargaining game of alternating offers between themselves. The argument for the uniqueness of the equilibrium outcome is similar to that in the proof of the first part of Proposition 3.5.

Now consider the case $\delta p_H^* < v_L$. The game has a stationary subgame perfect equilibrium in which the seller always proposes the price p^*, and accepts the highest proposed price when that price is at least v_L, trading with B_H if the proposed prices are equal. Both buyers propose the price v_L; B_H accepts any price at most equal to p^*, and B_L accepts any price less than v_L. Notice that the seller is better off accepting the price v_L than waiting to get the price p^* since $\delta p_H^* < v_L$ implies that $\delta p^* = \delta^2 v_L + \delta(1-\delta)v_H < v_L$, and B_H is indifferent between accepting p^* and waiting to get the price v_L, since $v_H - p^* = \delta(v_H - v_L)$.

We now show that in all subgame perfect equilibria the good is sold (to B_H if $v_H > v_L$) at the price p^*. Let M_s and m_s be the supremum and infimum, respectively, of the seller's payoff over all subgame perfect equilibria of the game in which the seller makes the first offer, and let M_I and m_I ($I = H, L$) be the supremum and infimum, respectively, of B_I's payoff over all subgame perfect equilibria of the game in which the buyers make the first offers.

Step 1. $m_H \geq v_H - \max\{v_L, \delta M_s\}$.

Proof. This follows from the facts that the seller must accept any price in excess of δM_s, and B_L never proposes a price in excess of v_L.

Step 2. $M_s \leq p^*$ $(= \delta v_L + (1-\delta)v_H)$.

Proof. We have $M_s \leq v_H - \delta m_H$ by the argument in the proof of Step 2 of Theorem 3.4, and thus by Step 1 we have $M_s \leq v_H - \delta(v_H - \max\{v_L, \delta M_s\})$. If $\delta M_s \leq v_L$ the result follows. If $\delta M_s > v_L$ then the result follows from the assumption that $\delta p_H^* < v_L$.

Step 3. $M_H \leq v_H - v_L$.

Proof. From Step 2 and $\delta p_H^* < v_L$ we have $\delta M_s < v_L$, so the seller must accept any price slightly less than v_L. If there is an equilibrium of the game in which the buyers make the first offers for which B_H's payoff exceeds $v_H - v_L$ then in this equilibrium B_L's payoff is 0, and hence B_L can profitably deviate by proposing a price close to v_L, which the seller accepts.

Step 4. $m_s \geq p^*$.

Proof. Since B_H must accept any price p for which $v_H - p > \delta M_H$, we have $m_s \geq v_H - \delta M_H \geq p^*$ (using Step 3).

We have now shown that $M_s = m_s = p^*$ and $M_H = m_H = v_H - v_L$. Since $p^* \geq v_L$, the sum of the payoffs of S and B_H is at least $p^* + \delta(v_H - v_L) = v_H$, so that the game must end with immediate agreement on the price p^*. If $v_H > v_L$ then $p^* > v_L$, so that it is B_H who accepts the first offer of the seller. □

Note that if $\delta = 1$ in this model then immediate agreement on any price between v_L and v_H is a subgame perfect equilibrium outcome. Note also that if B_L responds to an offer of the seller before rather than after B_H, or if the responses are simultaneous, then the result is the same.

9.4 Models with Choice of Partner

Here we study two models in which the seller chooses the buyer with whom to bargain. The models are related to those in Section 3.12; choosing to abandon one's current partner is akin to "opting out". In Section 3.12, the payoff to opting out is exogenous. Here, the corresponding payoff is determined by the outcome of the negotiations with the new buyer, which in turn is affected by the possibility that the seller can move back to the first buyer.

In both models, the seller and a buyer alternate offers until either one of them accepts an offer, or the seller abandons the buyer. In the latter case, the seller starts negotiating with the other buyer, until an offer is accepted or the seller returns to the first buyer. The main difference between the models lies in the times at which the seller may replace her partner. In the first model, the seller is the first to make an offer in any partnership, and can switch to the other buyer only at the beginning of a period in which she has to make an offer (cf. the model in Section 3.12.1). In the second model, it is the buyer who makes the first offer in any partnership, and the seller can switch to another buyer only at the beginning of a period in which the buyer has to make an offer (cf. the model in Section 3.12.2).

By comparison with the model of Section 9.3, the seller has an extra tool: she can threaten to terminate her negotiations with one of the buyers if he does not accept her demand. On the other hand, when matched with the seller a buyer is in a less competitive situation than in the model of public price announcements since he is the only buyer conversing with the seller.

9.4.1 The Case in Which the Seller Can Switch Partners Only Before Making an Offer

This model predicts a price equal to the equilibrium price in bilateral bargaining between the seller and B_H. The fact that the seller confronts more

than one buyer has no effect on the equilibrium price: the model does not capture any "competition" between the buyers.

Proposition 9.3 *In all subgame perfect equilibria the good is sold (to B_H if $v_H > v_L$) at the price $p_H^* = v_H/(1+\delta)$ (i.e. the unique subgame perfect equilibrium price of the bargaining game of alternating offers between the seller and B_H).*

Proof. We first describe a subgame perfect equilibrium with the properties given in the result. In this equilibrium, the seller always chooses B_H, proposes the price p_H^*, and accepts a price only if it is at least δp_H^*; buyer B_H proposes the price δp_H^*, and accepts any price at most equal to p_H^*; and buyer B_L proposes the price $\min\{v_L, \delta p_H^*\}$, and accepts any price at most equal to $\min\{v_L, p_H^*\}$.

We now prove that the payoff of the seller in all subgame perfect equilibria is p_H^*. Let M_s and m_s be the supremum and infimum, respectively, of the seller's payoff over all subgame perfect equilibria of the game in which the seller makes the first offer, and let M_I and m_I ($I = H,\ L$) be the suprema and infima, respectively, of B_I's payoff over all subgame perfect equilibria of the game in which B_I is bargaining with the seller and makes the first offer.

Step 1. $m_I \geq v_I - \delta M_s$ for $I = L,\ H$, $m_s \geq v_H - \delta M_H$, $M_s \leq \max_{I=L,H}(v_I - \delta m_I)$, and $M_H \leq v_H - \delta m_s$.

The proofs of these inequalities are very similar to the proofs of Steps 1 and 2 of the proof of Theorem 3.4.

Step 2. $M_s \leq p_H^*$.

Proof. By the first and third inequalities in Step 1 we have $M_s \leq \max_{I=L,H}(v_I - \delta(v_I - \delta M_s))$. Since $v_H - \delta(v_H - \delta M_s) \geq v_L - \delta(v_L - \delta M_s)$ for any value of M_s, we have $M_s \leq v_H/(1+\delta)$.

Step 3. $m_s \geq p_H^*$.

Proof. This follows from the second and fourth inequalities in Step 1.

From Steps 2 and 3 the seller's payoff in every subgame perfect equilibrium is precisely p_H^*. If $v_H > v_L$ then there is no equilibrium in which the seller trades with B_L, since in any such equilibrium the seller must obtain at least m_s and B_L must obtain at least δm_L, and $m_s + \delta m_L \geq v_H/(1+\delta) + \delta v_L - \delta^2 v_H/(1+\delta) = (1-\delta)v_H + \delta v_L > v_L$. Further, trade with B_H must occur in period 0 since $m_s + \delta m_H = v_H$. □

		H_1	H_2	L
S	proposes	p^*	p^*_H	p^*_H
	accepts	$p \geq \delta p^*$	$p \geq \delta p^*_H$	$p \geq \delta p^*_H$
	bargains with	B_H	B_H	B_L
B_I $(I = L, H)$	proposes	δp^*	δp^*_H	δp^*_H
	accepts	$p \leq p^*$	$p_1 \leq p^*_H$	$p \leq p^*_H$
	Transitions	Go to L if B_H rejects a price $p \leq p^*$.	Absorbing	Go to H_2 if B_L rejects a price $p \leq p^*_H$.

Table 9.2 A subgame perfect equilibrium for the model in Section 9.4.2 when p^*_H $(= v_H/(1+\delta)) < v_L$. The price p^* may take any value between p^*_H and v_H.

9.4.2 The Case in Which the Seller Can Switch Partners Only Before Her Partner Makes an Offer

In this case, the *buyer* makes the first offer when the seller switches partners. We restrict attention to the interesting case in which $p^*_H < v_L$. For any price p^* with $p^*_H \leq p^* \leq v_H$, Table 9.2 gives a subgame perfect equilibrium that ends with immediate agreement on δp^*. In any subgame starting in state H_1 the good is sold to B_H at the price p^* or δp^*, depending on who moves first; in any subgame starting in state H_2 the good is sold to B_H at the price p^*_H or δp^*_H; and in any subgame starting in state L the good is sold to B_L at the price p^*_H or δp^*_H.

To see that the strategy profile is a subgame perfect equilibrium notice the following. Once state H_2 is reached, the seller stays with B_H, and she and B_H behave as in the subgame perfect equilibrium of the game in which B_L is absent. In state H_1 buyer B_H prefers the price δp^* with one period of delay to the price p^* $(\delta(v_H - \delta p^*) > v_H - p^*)$. However, he is deterred from rejecting p^* by the transition to L, in which the good is sold to B_L. If $v_H > v_L$ then in state L buyer B_L prefers the price δp^*_H with one period of delay to the price p^*_H, but he is deterred from rejecting p^*_H by the transition to the absorbing state H_2. (If $v_H = v_L$ there is no need for deterrence, since B_L is indifferent between these two prices.)

Note that the game has other equilibria, some of which generate inefficient outcomes. For example, if the initial state is L then the strategy profile defined in the table is a subgame perfect equilibrium in which the good is sold to B_L.

We see that in a model in which the seller chooses whether or not to terminate bargaining with one buyer and move to the other, the results do not capture our intuition about competition between the buyers. In the first model, the presence of B_L is irrelevant for the equilibrium. The reason is clear, in light of the analysis in Section 3.12.1. The seller can never obtain more by moving to B_L than by staying with B_H, and thus a threat to move is not credible. In the second model, the ability to move to the other buyer after her offer is rejected enhances the power of the seller. In this case she can credibly threaten to abandon her current partner, and thus make a "take-it-or-leave-it" offer. This allows us to construct subgame perfect equilibria in which she obtains a price in excess of that which she would obtain in the absence of B_L.

9.5 A Model with More General Contracts and Resale

We conclude by investigating a model in which the range of contracts available to the agents is greater than it is in the models of the previous sections. The buyers are allowed to agree to a contract according to which B_H pays B_L a sum of money and in exchange B_L leaves the market. This contract leaves B_H alone in the market and thus presumably puts him in a better bargaining position. In addition, the seller and each of the buyers are allowed to agree to exchange the good for some sum of money, and if B_L buys the good then he is allowed to resell it to B_H. For simplicity we depart from the strategic approach of the previous sections and, as in Chapter 6, use the Nash solution to model bargaining; we restrict attention to the case $v_H > v_L$ and assume that $\delta = 1$.

The trading procedure is the following. In each period, two agents are matched and reach the agreement given by the Nash solution of the appropriate bargaining problem. If no agreement has been reached, then all three possible matches (including that between B_H and B_L) are equally probable. If the seller is matched with a buyer, then they agree on the amount of money paid by the buyer in exchange for the good. If the buyers are matched, then they agree on the amount of money B_H pays to B_L for him to leave the market.

We now consider the outcome after an agreement has been reached. If the seller reaches agreement with B_H, then the game ends. If the seller reaches agreement with B_L, then in the next period B_L is matched with B_H; the disagreement point gives B_L and B_H the payoffs v_L and 0, respectively (disagreement results in B_L consuming the good; B_L's payment is a sunk cost), and the size of the pie to be divided is v_H. Thus B_L and B_H agree on the price $v_L + (v_H - v_L)/2 = (v_H + v_L)/2$. If the two buyers reach agreement, then in the next period the seller is matched with B_H; the

disagreement point gives S and B_H each the payoff 0, and the size of the pie to be divided is v_H, so that S and B_H agree on the price $v_H/2$ (the payment to B_L is a sunk cost).

We now analyze the agreements reached in the first period. Denote by w_S, w_H, and w_L the expected payoffs of S, B_H, and B_L in the market. If the agents I and J who are matched fail to reach agreement, then the matching process occurs again in the next period. Thus the disagreement point for the bargaining in the first period is (w_I, w_J). Hence if S is matched with B_H in the first period then the Nash solution gives S the payoff $w_S + (v_H - w_S - w_H)/2$. If she is matched with B_L then the surplus to be divided is the price $(v_H + v_L)/2$ that B_L will obtain from B_H in the second period after he reaches agreement with S. Thus the Nash solution assigns her $w_S + [(v_H + v_L)/2 - w_S - w_L]/2$. If the two buyers are matched in the first period, then the surplus to be divided between B_H and the seller is v_H, so that the Nash solution assigns her $v_H/2$. Therefore

$$w_S = \tfrac{1}{3}\left(w_S + \frac{v_H - w_S - w_H}{2}\right) + \tfrac{1}{3}\left(w_S + \frac{(v_H + v_L)/2 - w_S - w_L}{2}\right) + \tfrac{1}{3}\left(\frac{v_H}{2}\right).$$

The first term corresponds to the case that S is matched first with B_H, the second to the case that S is matched first with B_L, and the third to the case that the two buyers are matched first. Similarly, we have

$$w_H = \tfrac{1}{3}\left(w_H + \frac{v_H - w_S - w_H}{2}\right) + \tfrac{1}{3}\left(\frac{v_H - v_L}{2}\right) + \tfrac{1}{3}\left(w_H + \frac{v_H/2 - w_H - w_L}{2}\right)$$

and

$$w_L = \tfrac{1}{3} \cdot 0 + \tfrac{1}{3}\left(w_L + \frac{(v_H + v_L)/2 - w_S - w_L}{2}\right) + \tfrac{1}{3}\left(w_L + \frac{v_H/2 - w_H - w_L}{2}\right).$$

The solution of this set of three equations is $(w_S, w_H, w_L) = (v_L/6 + v_H/2, v_H/2 - v_L/3, v_L/6)$.

An interesting feature of this vector of payoffs is its connection with the Shapley value. Recall that a cooperative game is specified by a function v that assigns to every coalition C its worth $v(C)$. In the market discussed here we have $v(S, B_H) = v(S, B_H, B_L) = v_H$, $v(S, B_L) = v_L$, and $v(C) = 0$ for all other coalitions C. The *Shapley value* of the cooperative game v assigns to each Player i the average, over all orderings of the players, of his marginal contribution $v(C \cup \{i\}) - v(C)$, where C is the set of players preceding i in the ordering. Thus in the market here the Shapley value assigns

$$\tfrac{1}{3} \cdot 0 + \tfrac{1}{6}v_L + \tfrac{1}{2}v_H = v_L/6 + v_H/2$$

to the seller,

$$\tfrac{1}{2} \cdot 0 + \tfrac{1}{3}(v_H - v_L) + \tfrac{1}{6}v_H = v_H/2 - v_L/3$$

to B_H, and $v_L/6$ to B_L. This vector is precisely the vector of payments that we isolated above. Note that the seller's payoff exceeds p_H^* (which is equal to $v_H/2$, since $\delta = 1$): the seller gains from the existence of B_L.

We have already mentioned that one of the attractions of models of matching and bargaining is that they enable us to interpret and better understand solution concepts from cooperative game theory. The model of this section illustrates this point.

Notes

The random matching model of Section 9.2 is based on Rubinstein and Wolinsky (1990); the proof of Proposition 9.1 is due to Shaked, and the nonstationary equilibrium for the case $v_H > v_L$ is due to Hendon and Tranæs (1991). The model in Section 9.3 is based on models of Binmore (1985) and Wilson (1984). The first model in Section 9.4 is due to Binmore (1985) and Wilson (1984); the second model is closely related to a model in Shaked (1994). Gul (1989) is the basis for the model of Section 9.5, although our interpretation is different from his.

A number of variations of the model in Section 9.2 have been investigated in the context of concrete economic problems. Among these is the model of Horn and Wolinsky (1988), in which the players are a firm and two unions. In this case the question whether an agreement between the firm and one of the unions is implemented immediately, or only after an agreement with the other union, is an important factor in determining the outcome. Related models are discussed by Davidson (1988), Jun (1989), and Fernandez and Glazer (1990). Bester (1988b) studies a model in which there is a single seller, who is randomly matched with a succession of buyers; the quality of the indivisible good that the seller holds is unknown to the buyers, and the reservation values of the buyers are unknown to the seller. Bester finds conditions under which there is an equilibrium in which price signals quality, and under which adverse selection leads a seller with a high-quality good to leave the market.

Gale (1988) and Peters (1991) study the relation between the equilibria of models in which, as in Section 9.3, sellers announce prices, which all buyers hear (*ex ante* pricing), and the equilibria of models in which (as in Section 9.2, for example) prices are determined by bargaining after a match is made (*ex post* pricing). Peters (1991) considers a model of a large market; when the agents' common discount factor is close to 1 the equilibrium sequence of *ex ante* prices as the market clears out approaches the competitive price. When demand and supply are relatively close, *ex ante* prices are lower than *ex post* prices; when excess demand is large, the reverse is true.

Shaked and Sutton (1984a) and Bester (1989a) study variations of the model in Section 9.4.1, in which the delay before the seller can make an offer to a *new* buyer may differ from the delay between any two successive periods of bargaining. (See also Muthoo (1993).) Shaked and Sutton use their model, in which a firm bargains with two workers, to study unemployment. Bester uses his model to replace the price-setting stage of Hotelling's model of spatial competition. Bester (1988a) is related; the aim is to explain the dependence of price on quality. Casella and Feinstein (1990, 1992) study a model in which the desire of a seller to move to a new buyer arises because inflation reduces the real value of the monetary holdings of her existing partner relative to that of a fresh buyer.

Peters (1988) studies a model that contains elements from the models of Sections 9.3 and 9.4. Sellers post prices, but a buyer who is matched with a seller has the option of making a counteroffer; the seller can accept this offer, reject it and continue bargaining, or terminate the match. When excess demand is small, posted prices are accepted in equilibrium; when it is large, they are not. The limit of the equilibrium outcome as the common discount factor approaches 1 is different from the competitive outcome. Peters (1989) studies a model in which, in each period, each seller chooses the trading rule she will use—i.e. the game that she will play with the buyer with whom she is matched. He shows that equilibrium trading rules lead to outcomes close to the competitive one.

The results of Gul (1989) are more general than those in Section 9.5. For a distinct but related implementation of the Shapley value, see Dow (1989). A steady-state model in which some agents are middlemen who buy from sellers and resell to buyers (and do not themselves consume the good) is studied by Rubinstein and Wolinsky (1987).

CHAPTER 10

The Role of Anonymity

10.1 Introduction

In this chapter we study the effect of the information structure on the relationship between market equilibria and competitive outcomes. As background for the analysis, recall that the models of Chapters 6 (Model B) and 8, in which all agents enter the market at once, yield competitive outcomes.

There are many aspects of the market about which an agent may or may not be informed. He may know the name of his opponent or may know only some of that agent's characteristics. He may remember his history in the market (whether he was matched, the characteristics of his opponent, the events in the match, etc.) or may retain only partial information about his experience. He may obtain information about the histories of other agents or may have no information at all about the events in bargaining sessions in which he did not take part.

In this chapter we focus on an assumption made in Chapter 8 that agents cannot condition their behavior in a bargaining encounter on their experience in previous encounters, or on the identity of their opponents. We refer to this as the "anonymity" assumption. We return to the model of Section 8.2. We change only the assumption about the agents' information;

we assume that they have full information about all past events. We show that under this assumption the outcome generated by a market equilibrium is not necessarily competitive.

10.2 The Model

For convenience we specify all the details of the model, although (as we noted above) the model is almost the same as that in Section 8.2. It is also closely related to the model of random matching studied in Section 9.2.

Goods A single indivisible good is traded for some quantity of a divisible good ("money").

Time Time is discrete and is indexed by the nonnegative integers.

Economic Agents In period 0, S identical sellers enter the market with one unit of the indivisible good each, and $B > S$ identical buyers enter with one unit of money each. No more agents enter at any later date. Each individual's preferences on lotteries over the pairs (p, t) giving the price and time at which a transaction is concluded satisfy the assumptions of von Neumann and Morgenstern. Each seller's preferences are represented by the utility function $\delta^t p$, where $0 < \delta \leq 1$, and each buyer's preferences are represented by the utility function $\delta^t(1 - p)$ (i.e. the reservation values of the seller and buyer are 0 and 1, respectively). If an agent never trades, then his utility is zero. In most of the chapter, we consider the case $\delta = 1$.

Matching In each period any remaining sellers and buyers are matched pairwise. The matching technology is such that each seller meets exactly one buyer and no buyer meets more than one seller in any period. Since there are fewer sellers than buyers, $B - S$ buyers are thus left unmatched in each period. The matching process is random: in each period all possible matches are equally probable, and the matching is independent across periods.

Bargaining After a buyer and a seller have been matched they engage in a short bargaining process. First, one of the matched agents is selected randomly (with probability 1/2) to propose a price between 0 and 1. Then the other agent responds by accepting the proposed price or rejecting it. Rejection dissolves the match, in which case the agents proceed to the next matching stage. If the proposal is accepted, the parties implement it and depart from the market.

What remains to be specified is the information structure. The natural case to consider seems to be that in which each agent fully recalls his own personal experience but does not have information about the events in matches in which he did not take part. However, to simplify the presentation we analyze a simpler case in which each agent does possess information about other matches.

Information In period t each agent has perfect information about all the events that occurred through period $t-1$, including the events in matches in which he did not participate. When taking an action in period t, however, each agent does *not* have any information about the other matches that are formed in that period or the actions that are taken by the members of those matches.

10.3 Market Equilibrium

In this section we show that the competitive outcome is not the unique market equilibrium outcome when an agent's information allows him to base his behavior on events that occurred in the past. If there is a single seller in the market, then in any given period at most one match is possible, so that the game is one of perfect information. In this case, we use the notion of subgame perfect equilibrium. When there is more than one seller the game is one of imperfect information, and we use the notion of sequential equilibrium.

A strategy for an agent in the game specifies an action (offer or response rule) in every period, for every history of the market up to the beginning of the period. For the sake of uniformity, we refer to a sequential equilibrium of the game as a *market equilibrium.*

Proposition 10.1 *If $\delta = 1$ then for every price p^* between 0 and 1, and for every one-to-one function β from the set of sellers to the set of buyers, there is a market equilibrium in which each seller s sells her unit of the good to buyer $\beta(s)$ for the price p^*.*

We give a proof only for the case $S = 1$, a case that reveals most of the ideas of the proof of the more general case. Before doing so, we give an intuitive description of an equilibrium with the properties claimed in the proposition.

The idea behind the equilibrium is that at any time a distinguished buyer has the "right" to purchase the seller's unit at the price p^*. If buyer i has the right, then in the equilibrium the seller offers buyer i, and no other buyer, the unit she owns at the price p^* and accepts an offer from buyer i, and from no one else, provided it is at least equal to p^*. Initially buyer $\beta(s)$

has the right to purchase the seller's unit at the price p^*, where s is the name of the seller. A buyer who has the right retains it unless one of the following events occurs.

1. The seller offers some other buyer, say i', a price in excess of p^*. In this event the right is transferred from the previous right-holder to i'.

2. A buyer who does not hold the right to purchase a unit at the price p^* proposes a price in excess of p^*. In this case no agent obtains or retains the right to purchase the good at the price of p^*; instead, the original right-holder obtains the right to purchase the good at the (unattractive) price of 1 (his reservation value).

Once some buyer has the right to purchase the good at the price of one, he retains this right whatever happens. Given the way in which the right to purchase the good is transferred, no buyer different from $\beta(s)$ has an incentive to offer a price in excess of p^* (for this will simply lead to the original right-holder obtaining the good at the price of one), and the seller has no incentive to offer the good to any buyer at a price in excess of p^* (for this will result in that buyer obtaining the right to buy the good at the price of p^*).

We turn now to a formal presentation of the equilibrium.

Proof of Proposition 10.1 for the case of a single seller. As usual, we describe each agent's strategy as an automaton. The states are $R(i)$ and $C(i)$ for $i = 1, 2, \ldots, B$. Their interpretations are as follows.

$R(i)$ Buyer i has the right to buy the unit from the seller at the price p^*.

$C(i)$ Buyer i has the right to buy the unit from the seller at the price 1.

The agents' actions and the transition rules between states *when the seller is matched with buyer i* are given in Table 10.1. The initial state is $R(\beta(s))$, and (as always) transitions between states take place immediately after the events that trigger them.

The outcome of the $(B + 1)$-tuple of strategies is the following. If the seller is matched with a buyer different from $\beta(s)$ and is chosen to make an offer, she proposes the price 1, so that the state remains $R(\beta(s))$, and the offer is rejected. If the seller is matched with a buyer different from $\beta(s)$ and the buyer is chosen to make an offer, then the buyer offers the price p^*, the state remains $R(\beta(s))$, and the seller rejects the offer. The first time that the seller is matched with buyer $\beta(s)$, the price p^* is proposed by whoever is chosen to make an offer, this proposal is accepted, the parties leave the market, and no further trade takes place.

		$R(i)$	$R(j)$, $j \neq i$	$C(i)$	$C(j)$, $j \neq i$
Seller	proposes	p^*	1	1	1
	accepts	$p \geq p^*$	$p = 1$	$p = 1$	no price
Buyer i	proposes	p^*	p^*	1	1
	accepts	$p \leq p^*$	$p \leq p^*$	$p \leq 1$	$p < 1$
	Transitions		Go to $R(i)$ if the seller proposes p with $p^* < p < 1$.	Absorbing	Absorbing
			Go to $C(j)$ if Buyer i proposes p with $p^* < p < 1$.		

Table 10.1 The agents' actions and the transitions between states when the seller is matched with buyer i.

To see that the strategy profile is a subgame perfect equilibrium, suppose that the current state is $R(h)$, and consider two deviations that might upset the seller's "plan" to sell her good to buyer h. First, suppose that the seller offers a price in excess of p^* to a different buyer, say i. In this case the state changes to $R(i)$, and the buyer rejects the offer. It is optimal for the buyer to behave in this way since, given the state is $R(i)$, the strategies lead to his eventually receiving the good at the price p^*. Thus the seller does not gain from this deviation.

Second, suppose that buyer i, with $i \neq h$, proposes a price in excess of p^*, but less than 1. Then the state changes to $C(h)$, and the seller rejects the offer. It is optimal for the seller to act in this way because, starting from state $C(h)$, the strategies lead to the seller obtaining the price 1 from buyer h. Given this, buyer i does not benefit from the deviation.

Finally, if the current state is $C(i)$, it never changes; the good is eventually sold to buyer i at the price of 1, and no deviation can make any agent better off. □

Notice that a buyer's personal history is not sufficient for him to calculate the state. For example, if buyer $\beta(s)$ is not matched in the first period, then he needs to know what happened in that period in order to calculate the state in the second period. However, one can construct equilibria with the same outcome as the one here, in which each agent bases his behavior only on his own history. (See Rubinstein and Wolinsky (1990) for details.)

We can simplify the equilibrium given in the proof by replacing all the states $C(i)$ by a single state C, in which the seller offers and accepts the competitive price from *any* buyer with whom he is matched, and all the buyers accept and offer the price of one. However, this equilibrium is not robust to the following modification of the model. Suppose that the set of possible prices is discrete, and does not include 1. Then the competitive price is the largest price less than 1, and all buyers prefer to obtain the good at this price to not trading at all. Suppose that a buyer who does not have the right to purchase the good at the price p^* deviates from the strategy described in the proof by offering a price in excess of p^*. Then the state becomes C, in which there is a positive probability that this buyer obtains the good at the competitive price. Thus the buyer benefits from his deviation, and the strategy profile is no longer a subgame perfect equilibrium.

The model has a great multiplicity of equilibria. The proposition shows that all prices p^* can be sustained in market equilibria. Further, for each price p^* there is a rich set of market equilibria (in addition to that described in the proof of the proposition) supporting that price. The interest of the model derives from the character of the equilibrium we constructed in the proof. This equilibrium is interesting because it captures a social institution that is close (but not identical) to some that we observe. For example, the workers in a firm may have the right to buy that firm at a certain price; a neighbor may have priority in buying a piece of land; and Academic Press has the right to buy any book on bargaining that we write. Although there are many equilibria in which all units of the good are sold at the price p^*, and although some of them are more simply stated, we have chosen one equilibrium because of its attractive interpretation. In any given context, the appeal of the equilibrium we describe depends on how natural are the price p^* and the identity of the buyer $\beta(s)$. The price p^* may be determined, for example, by considerations of fairness, and the identity of the right-holder may be an expression of a social arrangement that gives special priority to a particular potential buyer. The existence of such an explanation of the price p^* and the asymmetric statuses of the buyers is necessary for the result to be of interest.

In the equilibria shown to exist by the proposition, all trades occur at the same price. However, there are other equilibria in which different prices are obtained by different sellers. For example, consider the case of two sellers and two buyers. Let p_1 and p_2 be two different prices. The following is a market equilibrium in which seller i sells the good to buyer i at the price p_i, $i = 1, 2$. Seller i offers buyer i the price p_i and accepts from buyer i any price of p_i or more. She offers buyer j the price 1 and rejects any price below 1 that buyer j offers. Analogously buyer i offers seller i the

price p_i and accepts from buyer i any price of p_i or less. He offers seller j the price 0 and rejects any price above 0 offered by seller j. If one of the sellers deviates, then the agents continue with the equilibrium strategies described in the proposition for the uniform price of 0, while if one of the buyers deviates, then the agents continue with the equilibrium strategies described in the proposition for the uniform price of 1.

Note that the strategy profile constructed in the proof of the proposition is not a market equilibrium when the market contains a single seller and two buyers b_H and b_L with reservation values $v_H > v_L > 0$, and the set of possible prices is discrete, includes a price between v_L and v_H, and does not include v_H. Obviously p^* cannot exceed v_L. If $p^* \leq v_L$ the strategy profile is not a market equilibrium for the following reason. When b_L holds the right to purchase the good at the price p^*, the seller must reject any offer by b_H that is above v_L. Therefore the price at which the good is sold in $C(b_L)$ must exceed v_L. But if the price attached to $C(b_L)$ exceeds v_L, then it is not optimal for b_L to purchase the good at this price. We know of no result that characterizes the set of market equilibria in this case.

10.4 The No-Discount Assumption

The assumption that the agents are indifferent to the timing of their payoffs is crucial to the proof of Proposition 10.1. Under this assumption, an agent is content to wait as long as necessary to be matched with the "right" partner. If he discounts future payoffs, then he prefers to trade at any given price as soon as possible, and the equilibrium of Proposition 10.1 disintegrates. In this case the model, for $S = 1$, is the same as that in Section 9.2.1. We showed there that there is a unique market equilibrium in which all transactions are concluded in the first period. (Proposition 9.1 covers only the case $B = 2$, but the extension is immediate.) In this equilibrium the seller always proposes the price $p_s(B)$, each buyer always proposes the price $p_b(B)$, and these prices are always accepted. The prices satisfy the following pair of equations.

$$\begin{aligned} p_b(B) &= \delta(p_s(B) + p_b(B))/2 \\ 1 - p_s(B) &= \delta(1 - p_s(B) + 1 - p_b(B))/2B. \end{aligned}$$

For $B > 1$ the limit as $\delta \to 1$ of both prices is the competitive price of 1. For $B = 1$ the equations define the unique subgame perfect equilibrium in a bargaining game of alternating offers in which the proposer is chosen randomly at the beginning of each period (see Section 3.10.3). The limit as $\delta \to 1$ of both agreed-upon prices $p_s(1)$ and $p_b(1)$ in this case is 1/2.

This result, especially for the case $\delta \to 1$, seems at first glance to cast doubt on the significance of Proposition 10.1. We argue that upon closer

examination the assumption that agents discount future payoffs, when combined with the other assumptions of the model, is not as natural as it seems. The fact that agents discount the future not only makes a delay in reaching agreement costly; the key fact in this model is that it makes holding a special relationship costly. A buyer and a seller who are matched are forced to separate at the end of the bargaining session even if they have a special "personal relationship". The chance that they will be reunited is the same as the chance that each of them will meet another buyer or seller. Thus there is a "tax" on personal relationships, a tax that prevents the formation of such relationships in equilibrium. It seems that this tax does not capture any realistic feature of the situations we observe.

We now try to separate the two different roles that discounting plays in the model. Remove the assumption that pairs have to separate at the end of a bargaining session; assume instead that each partner may stay with his current partner for another period or return to the pool of agents waiting to be matched in the next period. Suppose that the agents make the decision whether or not to stay with their current partner simultaneously. These assumptions do not penalize personal relationships, and indeed the results show that noncompetitive prices are consistent with subgame perfect equilibrium.

The model is very similar to that of Section 9.4.2. Here the proposer is selected randomly, and the seller may switch buyers at the beginning of each period. In the model of Section 9.4.2 the agents take turns in making proposals and the seller may switch buyers only at the beginning of a period in which her partner is scheduled to make an offer. The important feature of the model here that makes it similar to that of Section 9.4.2 rather than that of Section 9.4.1 is that the seller is allowed to leave her partner after he rejects her offer, which, as we saw, allows the seller to make what is effectively a "take-it-or-leave-it" offer.

As in Section 9.4.2 we can construct subgame perfect equilibria that support a wide range of prices. Suppose for simplicity that there is a single seller (and an arbitrary number B of buyers). For every p_s^* such that $p_s(1) \leq p_s^* \leq p_s(B)$ we can construct a subgame perfect equilibrium in which immediate agreement is reached on either the price p_s^*, or the price p_b^* satisfying $p_b^* = \delta(p_s^* + p_b^*)/2$, depending on the selection of the first proposer. In this equilibrium the seller always proposes p_s^*, accepts any price of p_b^* or more, and stays with her partner unless he rejected a price of at most p_s^*. Each buyer proposes p_b^*, accepts any price of p_s^* or less, and never abandons the seller.

Recall that $p_s(1)$ (which depends on δ) is the offer made by the seller in the unique subgame perfect equilibrium of the game in which there is a single buyer; $p_s(B)$ is the offer made by the seller when there are B buyers

and partners are forced to separate at the end of each period. The limits of $p_s(1)$ and $p_s(B)$ as δ converges to 1 are 1/2 and 1, respectively. Thus when δ is close to 1 almost all prices between 1/2 and 1 can be supported as subgame perfect equilibrium prices.

Thus when partners are not forced to separate at the end of each period, a wide range of outcomes—not just the competitive one—can be supported by market equilibria even if agents discount the future. We do not claim that the model in this section is a good model of a market. Moreover, the set of outcomes predicted by the theory includes the competitive one; we have not ruled out the possibility that another theory will isolate the competitive outcome. However, we have shown that the fact that agents are impatient does not automatically rule out noncompetitive outcomes when the other elements of the model do not unduly penalize "personal relationships".

10.5 Market Equilibrium and Competitive Equilibrium

"Anonymity" is sometimes stated as a condition that must be satisfied in order for an application of a competitive model to be reasonable. We have explored the meaning of anonymity in a model in which agents meet and bargain over the terms of trade. As Proposition 8.2 shows, when agents are anonymous, the only market equilibrium is competitive. When agents have sufficiently detailed information about events that occurred in the past and recognize their partners, then noncompetitive outcomes can emerge, even though the matching process is anonymous (agents are matched randomly).

The fact that this result is sensitive to our assumption that there is no discounting can be attributed to other elements of the model, which inhibit the agents' abilities to form special relationships. In our models, matches are random, and partners are forced to separate at the end of each period. If the latter assumption is modified, then we find that once again special relationships can emerge, and noncompetitive outcomes are possible.

We do not have a theory to explain how agents form special relationships. But the results in this chapter suggest that there is room for such a theory in any market where agents are not anonymous.

Notes

This chapter is based on Rubinstein and Wolinsky (1990).

References

The numbers in brackets after each reference are the page numbers on which the reference is cited. The hyperlinks lead to reviews of the items on the American Mathematical Society's MathSciNet. Depending on the services to which your institution subscribes, the page containing a review may contain also a link that allows you to check the availability of the item in your institution's library.

Admati, A. R., and M. Perry (1987), "Strategic Delay in Bargaining", *Review of Economic Studies* **54**, 345–364. [119]

Admati, A. R. and M. Perry (1991), "Joint Projects without Commitment", *Review of Economic Studies* **58**, 259–276. [67]

Anbarci, N. (1993), "Noncooperative Foundations of the Area Monotonic Solution", *Quarterly Journal of Economics* **108**, 245–258. [90]

Aumann, R. J. (1959), "Acceptable Points in General Cooperative n-Person Games", pp. 287–324 in A. W. Tucker and R. D. Luce (eds.), *Contributions to the Theory of Games, Vol. IV*, Princeton University Press. [65]

Ausubel, L. M., and R. J. Deneckere (1989a), "Reputation in Bargaining and Durable Goods Monopoly", *Econometrica* **57**, 511–531. [106]

Ausubel, L. M. and R. J. Deneckere (1992a), "Durable Goods Monopoly with Incomplete Information", *Review of Economic Studies* **59**, 187–203. [119]

Ausubel, L. M. and R. J. Deneckere (1992b), "Bargaining and the Right to Remain Silent", *Econometrica* **60**, 597–625. [119]

Baron, D. P., and J. A. Ferejohn (1987), "Bargaining and Agenda Formation in Legislatures", *American Economic Review* **77** (Papers and Proceedings), 303–309. [67]

Baron, D. P., and J. A. Ferejohn (1989), "Bargaining in Legislatures", *American Political Science Review* **83**, 1181–1206. [67]

Bester, H. (1988a), "Bargaining, Search Costs and Equilibrium Price Distributions", *Review of Economic Studies* **55**, 201–214. [188]

Bester, H. (1988b), "Qualitative Uncertainty in a Market with Bilateral Trading", *Scandinavian Journal of Economics* **90**, 415–434. [187]

Bester, H. (1989a), "Noncooperative Bargaining and Spatial Competition", *Econometrica* **57**, 97–113. [188]

Bester, H. (1989b), "Non-Cooperative Bargaining and Imperfect Competition: A Survey", *Zeitschrift für Wirtschafts- und Sozialwissenschaften* **109**, 265–286. [6]

Bikhchandani, S. (1986), "A Bargaining Model with One-Sided Incomplete Information about Reservation Prices", unpublished paper, Graduate School of Management, University of California, Los Angeles. [119]

Bikhchandani, S. (1992), "A Bargaining Model with Incomplete Information", *Review of Economic Studies* **59**, 187–203. [119]

Binmore, K. G. (1985), "Bargaining and Coalitions", pp. 269–304 in Roth (1985). [65, 187]

Binmore, K. G. (1987a), "Nash Bargaining Theory II", pp. 61–76 in Binmore and Dasgupta (1987). [65, 89, 90]

Binmore, K. G. (1987b), "Perfect Equilibria in Bargaining Models", pp. 77–105 in Binmore and Dasgupta (1987). [54, 66]

Binmore, K. G. (1987c), "Nash Bargaining and Incomplete Information", pp. 155–192 in Binmore and Dasgupta (1987). [90]

Binmore, K. G., and P. Dasgupta (1987), *The Economics of Bargaining*, Oxford: Blackwell.

Binmore, K. G., and M. J. Herrero (1988a), "Matching and Bargaining in Dynamic Markets", *Review of Economic Studies* **55**, 17–31. [136, 170, 171]

Binmore, K. G., and M. J. Herrero (1988b), "Security Equilibrium", *Review of Economic Studies* **55**, 33–48. [148]

Binmore, K. G., M. J. Osborne, and A. Rubinstein (1992), "Noncooperative Models of Bargaining", pp. 179–225 in R. J. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications* (Volume 1), Amsterdam: North-Holland. [6]

Binmore, K. G., A. Rubinstein, and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modelling", *Rand Journal of Economics* **17**, 176–188. [90]

Binmore, K. G., A. Shaked, and J. Sutton (1989), "An Outside Option Experiment", *Quarterly Journal of Economics* **104**, 753–770. [65]

Bulow, J., and K. Rogoff (1989), "A Constant Recontracting Model of Sovereign Debt", *Journal of Political Economy* **97**, 155–178. [67]

Butters, G. R. (1977), "Equilibrium Price Distributions in a Random Meetings Market", unpublished paper, Princeton University. [136]

Carlsson, H. (1991), "A Bargaining Model where Parties Make Errors", *Econometrica* **59**, 1487–1496. [90]

Casella, A., and J. S. Feinstein (1990), "Economic Exchange during Hyperinflation", *Journal of Political Economy* **98**, 1–27. [188]

Casella, A., and J. S. Feinstein (1992), "A Note on Bargaining and Inflation", *Economics Letters* **38**, 393–398. [188]

Chae, S., and J.-A. Yang (1988), "The Unique Perfect Equilibrium of an N-Person Bargaining Game", *Economics Letters* **28**, 221–223. [67]

Chatterjee, K., B. Dutta, D. Ray, and K. Sengupta (1993), "A Non-Cooperative Theory of Coalitional Bargaining", *Review of Economic Studies* **60**, 463–477. [67]

Chatterjee, K., and L. Samuelson (1987), "Bargaining with Two-sided Incomplete Information: An Infinite Horizon Model with Alternating Offers", *Review of Economic Studies* **54**, 175–192. [120]

Chatterjee, K., and L. Samuelson (1988), "Bargaining under Two-Sided Incomplete Information: The Unrestricted Offers Case", *Operations Research* **36**, 605–618. [119]

Chatterjee, K. and L. Samuelson (1990), "Perfect Equilibria in Simultaneous-Offers Bargaining", *International Journal of Game Theory* **19**, 237–267. [67]

Chikte, S. D. and S. D. Deshmukh (1987), "The Role of External Search in Bilateral Bargaining', *Operations Research* **35**, 198–205. [67]

Cho, I.-K. (1989), "Characterization of Stationary Equilibria in Bargaining Models with Incomplete Information", unpublished paper, Department of Economics, University of Chicago. [119]

Cho, I.-K., and D. M. Kreps (1987), "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics* **102**, 179–221. [107, 112]

Clemhout, S., and H. Y. Wan, Jr. (1988), "A General Dynamic Model of Bargaining—The Perfect Information Case", pp. 293–305 in *Advances in Optimization and Control* (H. A. Eiselt and G. Pederzoli, eds.), Springer-Verlag, Berlin. [67]

Cothren, R., and M. A. Loewenstein (n.d.), "Quality Signals and Asymmetric Information in a Sequential Bargaining Game", unpublished

paper, Virginia Polytechnic Institute and State University. [119]
Cramton, P. C. (1992), "Strategic Delay in Bargaining with Two-Sided Uncertainty', *Review of Economic Studies* **59**, 205–225. [119]
Dasgupta, P., and E. S. Maskin (1989), "Bargaining and Destructive Power", Discussion Paper 1432, Harvard Institute of Economic Research, Harvard University. [90]
Davidson, C. (1988), "Multiunit Bargaining in Oligopolistic Industries", *Journal of Labor Economics* **6**, 397–422. [187]
Derman, C. (1970), *Finite State Markovian Decision Processes*, New York: Academic Press. [44, 146]
Diamond, P. A. (1981), "Mobility Costs, Frictional Unemployment, and Efficiency", *Journal of Political Economy* **89**, 798–812. [136]
Diamond, P. A., and E. Maskin (1979), "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States", *Bell Journal of Economics* **10**, 282–316. [136]
Dow, G. K. (1989), "Knowledge Is Power: Informational Precommitment in the Capitalist Firm", *European Journal of Political Economy* **5**. [188]
Dutta, B., and L. Gevers (1984), "On Majority Rules, Veto Rights and Perfect Equilibrium Allocations of a Shrinking Cake", Cahiers de la Faculté des Sciences Economiques et Sociales de Namur, Série Recherche, 60, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium. [67]
Fernandez, R. and J. Glazer (1990), "The Scope for Collusive Behavior among Debtor Countries", *Journal of Development Economics* **32**, 297–313. [187]
Fernandez, R. and J. Glazer (1991), "Striking for a Bargain between Two Completely Informed Agents", *American Economic Review* **81**, 240–252. [66]
Fernandez, R. and R. W. Rosenthal (1990), "Strategic Models of Sovereign-Debt Renegotiations", *Review of Economic Studies* **57**, 331–349. [67]
Fershtman, C. (1989), "Simultaneous Moves Multi-Person Continuous Time Concession Game", *Theory and Decision* **26**, 81–90. [67]
Fershtman, C. (1990), "The Importance of the Agenda in Bargaining", *Games and Economic Behavior* **2**, 224–238. [67]
Fishburn, P. C., and A. Rubinstein (1982), "Time Preference", *International Economic Review* **23**, 677–694. [33, 34, 83]
Fudenberg, D., D. Levine, and J. Tirole (1985), "Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information", pp. 73–98 in Roth (1985). [120]
Gale, D. (1986a), "Bargaining and Competition Part I: Characterization",

Econometrica **54**, 785–806. [160, 170]

Gale, D. (1986b), "Bargaining and Competition Part II: Existence", *Econometrica* **54**, 807–818. [168, 170]

Gale, D. (1986c), "A Simple Characterization of Bargaining Equilibrium in a Large Market Without the Assumption of Dispersed Characteristics", Working Paper 86-05, Center for Analytic Research in Economics and the Social Sciences, University of Pennsylvania. [158, 170]

Gale, D. (1986d), "A Strategic Model of Trade with Money as a Medium of Exchange", Working Paper 86-04, Center for Analytic Research in Economics and the Social Sciences, University of Pennsylvania. [149]

Gale, D. (1986e), "A Strategic Model of Labor Markets with Incomplete Information", unpublished paper, University of Pittsburgh. [171]

Gale, D. (1987), "Limit Theorems for Markets with Sequential Bargaining", *Journal of Economic Theory* **43**, 20–54. [136, 147, 170]

Gale, D. (1988), "Price Setting and Competition in a Simple Duopoly Model", *Quarterly Journal of Economics* **103**, 729–739. [187]

Green, E. J. (1992), "Eliciting Traders' Knowledge in 'Frictionless' Asset Market", pp. 332–355 in *Game theory and economic applications (New Delhi, 1990)*, Lecture Notes in Economic and Mathematical Systems, Vol. **389**, Springer, Berlin. [148]

Grossman, S. J., and M. Perry (1986), "Sequential Bargaining under Asymmetric Information", *Journal of Economic Theory* **39**, 120–154. [107, 119]

Gul, F. (1989), "Bargaining Foundations of Shapley Value", *Econometrica* **57**, 81–95. [187, 188]

Gul, F., and H. Sonnenschein (1988), "On Delay in Bargaining with One-Sided Uncertainty", *Econometrica* **56**, 601–611. [105, 119]

Gul, F., H. Sonnenschein, and R. Wilson (1986), "Foundations of Dynamic Monopoly and the Coase Conjecture", *Journal of Economic Theory* **39**, 155–190. [106]

Haller, H. (1986), "Non-Cooperative Bargaining of $N \geq 3$ Players", *Economics Letters* **22**, 11–13. [67]

Haller, H. (1991), "Wage Bargaining as a Strategic Game", pp. 230–241 in R. Selten (ed.), *Game Equilibrium Models III: Strategic Bargaining*, Berlin: Springer-Verlag. [66]

Haller, H., and S. Holden (1990), "A Letter to the Editor on Wage Bargaining", *Journal of Economic Theory* **52**, 232–236. [66]

Harrington, Jr., J. E. (1990), "The Role of Risk Preferences in Bargaining when Acceptance of a Proposal Requires Less than Unanimous

Approval", *Journal of Risk and Uncertainty* **3**, 135–154. [67]

Harsanyi, J. C. (1967/8), "Games with Incomplete Information Played by 'Bayesian' Players", Parts I, II, and III, *Management Science* **14**, 159–182, 320–334, 486–502. [92]

Harsanyi, J. C. (1974), "An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition", *Management Science (Theory Series)* **20**, 1472–1495. [67]

Harsanyi, J. C. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press. [6]

Harsanyi, J. C. (1981), "The Shapley Value and the Risk-Dominance Solutions of Two Bargaining Models for Characteristic-Function Games", pp. 43–68 in R. J. Aumann, J. C. Harsanyi, W. Hildenbrand, M. Maschler, M. A. Perles, J. Rosenmüller, R. Selten, M. Shubik, and G. L. Thompson, *Essays in Game Theory and Mathematical Economics*, Mannheim: Bibliographisches Institut. [67]

Harsanyi, J. C., and R. Selten (1972), "A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information", *Management Science* **18**, P-80–P-106. [27, 119]

Hart, S. (1979), "Lecture Notes: Special Topics in Game Theory", unpublished paper, Institute for Mathematical Studies in the Social Sciences, Stanford University. [27]

Hendon, E., and T. Tranæs (1991), "Sequential Bargaining in a Market with One Seller and Two Different Buyers', *Games and Economic Behavior* **3**, 453–466. [187]

Herrero, M. J. (1984), "Bargaining and Involuntary Unemployment", unpublished paper, London School of Economics. [65]

Herrero, M. J. (1988), "Single-Package versus Issue-by-Issue Bargaining", unpublished paper, Carnegie-Mellon University. [67]

Herrero, M. J. (1989), "The Nash Program: Non-convex Bargaining Problems", *Journal of Economic Theory* **49**, 266–277. [90]

Hopcroft, J. E., and J. D. Ullman (1979), *Introduction to Automata Theory, Languages, and Computation*, Reading, Massachusetts: Addison-Wesley. [40]

Horn, H., and A. Wolinsky (1988), "Worker Substitutability and Patterns of Unionisation", *Economic Journal* **98**, 484–497. [187]

Howard, J. V. (1992), "A Social Choice Rule and Its Implementation in Perfect Equilibrium", *Journal of Economic Theory* **56**, 142–159. [90]

Jones, S. R. G., and C. J. McKenna (1988), "Inventories, Strike Funds and Bargaining Outcomes", Discussion Paper 88-17, Department of

Economics, University of British Columbia. [66]

Jun, B. H. (1987), "A Strategic Model of 3-Person Bargaining", unpublished paper, State University of New York at Stony Brook. [67]

Jun, B. H. (1989), "Non-cooperative Bargaining and Union Formation", *Review of Economic Studies* **56**, 59–76. [187]

Kalai, E. (1977), "Nonsymmetric Nash Solutions and Replications of 2-Person Bargaining", *International Journal of Game Theory* **6**, 129–133. [27]

Kalai, E. (1985), "Solutions to the Bargaining Problem", pp. 77–105 in L. Hurwicz, D. Schmeidler, and H. Sonnenschein (eds.), *Social Goals and Social Organization*, Cambridge University Press. [27]

Kalai, E., and M. Smorodinsky (1975), "Other Solutions to Nash's Bargaining Problem", *Econometrica* **43**, 513–518. [27]

Kihlstrom, R. E., A. E. Roth, and D. Schmeidler (1981), "Risk Aversion and Solutions to Nash's Bargaining Problem", pp. 65–71 in O. Moeschlin and D. Pallaschke (eds.), *Game Theory and Mathematical Economics*, Amsterdam: North-Holland. [26]

Kiyotaki, N., and R. Wright (1989), "On Money as a Medium of Exchange", *Journal of Political Economy* **97**, 927–954. [149]

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971), *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*, New York: Academic Press. [23]

Krelle, W. (1975), "A New Theory of Bargaining", Working Paper 70, Institut für Gesellschafts- und Wirtschaftswissenschaften, Universität Bonn. [66]

Krelle, W. (1976), *Preistheorie*, Volume II, Tübingen: J. C. B. Mohr. [66]

Kreps, D. M. (1990), *A Course in Microeconomic Theory*, Princeton: Princeton University Press. [x]

Kreps, D. M., and G. Ramey (1987), "Structural Consistency, Consistency, and Sequential Rationality", *Econometrica* **55**, 1331–1348. [95]

Kreps, D. M., and R. Wilson (1982), "Sequential Equilibria", *Econometrica* **50**, 863–894. [92, 95]

Leitmann, G. (1973), "Collective Bargaining: A Differential Game", *Journal of Optimization Theory and Application* **11**, 405–412. [67]

Luce, R. D., and H. Raiffa (1957), *Games and Decisions*, New York: Wiley. [x]

Madrigal, V., T. C. C. Tan, and S. Ribeiro da Costa Werlang (1987), "Support Restrictions and Sequential Equilibria", *Journal of Economic Theory* **43**, 329–334. [96]

Matsuo, T. (1989), "On Incentive Compatible, Individually Rational, and Ex Post Efficient Mechanisms for Bilateral Trading", *Journal of Eco-*

nomic Theory **49**, 189–194. [119]

McDonald, I. M., and R. M. Solow (1981), "Wage Bargaining and Employment", *American Economic Review* **71**, 896–908. [27]

McLennan, A. (1988), "Bargaining between Two Symmetrically Informed Agents", unpublished paper, University of Minnesota. [90]

McLennan, A. and H. Sonnenschein (1991), "Sequential Bargaining as a Noncooperative Foundation for Walrasian Equilibrium", *Econometrica* **59**, 1395–1424. [168, 171]

Mortensen, D. T. (1982a), "Property Rights and Efficiency in Mating, Racing, and Related Games", *American Economic Review* **72**, 968–979. [136]

Mortensen, D. T. (1982b), "The Matching Process as a Noncooperative Bargaining Game", pp. 233–254 in J. J. McCall (ed.), *The Economics of Information and Uncertainty*, Chicago: University of Chicago Press. [136]

Moulin, H. (1984), "Implementing the Kalai-Smorodinsky Bargaining Solution", *Journal of Economic Theory* **33**, 32–45. [90]

Muthoo, A. (1989), "A Note on the Strategic Role of Outside Options in Bilateral Bargaining", unpublished paper, Department of Economics, London School of Economics. [67]

Muthoo, A. (1990), "Bargaining without Commitment", *Games and Economic Behavior* **2**, 291–297. [66]

Muthoo, A. (1991), "A Note on Bargaining Over a Finite Number of Feasible Agreements", *Economic Theory* **1**, 290–292. [66]

Muthoo, A. (1992), "Revocable Commitment and Sequential Bargaining", *Economic Journal* **102**, 378–387. [66]

Muthoo, A. (1993), "Sequential Bargaining and Competition", *Economic Theory* **3**, 353–363. [188]

Myerson, R. B., and M. A. Satterthwaite (1983), "Efficient Mechanisms for Bilateral Trading", *Journal of Economic Theory* **29**, 265–281. [119]

Nash, J. F. (1950a), "The Bargaining Problem", *Econometrica* **18**, 155–162. [3, 9, 26]

Nash, J. F. (1950b), "Equilibrium Points in N-Person Games", *Proceedings of the National Academy of Sciences* (U. S. A.) **36**, 48–49. [41]

Nash, J. F. (1951), "Non-Cooperative Games", *Annals of Mathematics* **54**, 286–295. [41]

Nash, J. F. (1953), "Two-Person Cooperative Games", *Econometrica* **21**, 128–140. [11, 26, 27, 29, 70, 76, 78, 89, 90]

Okada, A. (1988b), "A Noncooperative Bargaining Model for the Core in n-Person Characteristic Function Games", unpublished paper, Department of Information Sciences, Tokyo Institute of Technology.

[67]

Okada, A. (1991a), "A Two-Person Repeated Game with Long-Term Contracts", pp. 34–47 in R. Selten (ed.), *Game Equilibrium Models III: Strategic Bargaining*, Berlin: Springer-Verlag. [66]

Okada, A. (1991b), "A Noncooperative Approach to the Nash Bargaining Problem", pp. 7–33 in R. Selten (ed.), *Game Equilibrium Models III: Strategic Bargaining*, Berlin: Springer-Verlag. [66]

Osborne, M. J. (1985), "The Role of Risk Aversion in a Simple Bargaining Model", pp. 181–213 in Roth (1985). [120]

Owen, G. (1982), *Game Theory* (2nd. edition), New York: Academic Press. [27]

Perry, M. (1986), "An Example of Price Formation in Bilateral Situations: A Bargaining Model with Incomplete Information", *Econometrica* **54**, 313–321. [119]

Perry, M. and P. J. Reny (1993), "A Non-cooperative Bargaining Model with Strategically Timed Offers", *Journal of Economic Theory* **59**, 50–77. [66]

Peters, M. (1988), "Ex Ante Pricing and Bargaining", unpublished paper, University of Toronto. [188]

Peters, M. (1989), "Stable Pricing Institutions are Walrasian", unpublished paper, University of Toronto. [188]

Peters, M. (1991), "Ex Ante Price Offers in Matching Games: Non-Steady States", *Econometrica* **59**, 1425–1454. [187]

Ponsati-Obiols, C. (1989), "Two-Sided Incomplete Information Bargaining with a Finite Set of Possible Agreements", Economics Discussion Paper 57, Bellcore (Morristown, New Jersey). [120]

Ponsati-Obiols, C. (1992), "Unique Equilibrium in a Model of Bargaining over Many Issues", *Annales d'Economie et de Statistique* **25–26**, 81–100. [120]

Roth, A. E. (1977), "Individual Rationality and Nash's Solution to the Bargaining Problem", *Mathematics of Operations Research* **2**, 64–65. [27]

Roth, A. E. (1979), *Axiomatic Models of Bargaining*, Berlin: Springer-Verlag. [27]

Roth, A. E. (1985), *Game-Theoretic Models of Bargaining*, Cambridge University Press. [119]

Roth, A. E. (1988), "Laboratory Experimentation in Economics: A Methodological Overview", *Economic Journal* **98**, 974–1031. [6]

Roth, A. E. (1989), "Risk Aversion and the Relationship between Nash's Solution and Subgame Perfect Equilibrium of Sequential Bargaining", *Journal of Risk and Uncertainty* **2**, 353–365. [90]

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model",

Econometrica **50**, 97–109. [49, 65]

Rubinstein, A. (1985a), "A Bargaining Model with Incomplete Information about Time Preferences", *Econometrica* **53**, 1151–1172. [118, 119]

Rubinstein, A. (1985b), "Choice of Conjectures in a Bargaining Game with Incomplete Information", pp. 99–114 in Roth (1985). [99, 118, 119]

Rubinstein, A. (1986), "Finite Automata Play the Repeated Prisoner's Dilemma", *Journal of Economic Theory* **39**, 83–96. [40]

Rubinstein, A. (1987), "A Sequential Strategic Theory of Bargaining", pp. 197–224 in T. F. Bewley (ed.), *Advances in Economic Theory*, Cambridge University Press. [65]

Rubinstein, A. (1989), "Competitive Equilibrium in a Market with Decentralized Trade and Strategic Behavior: An Introduction", pp. 243–259 in G. R. Feiwel (ed.), *The Economics of Imperfect Competition and Employment*, Basingstoke: Macmillan. [136]

Rubinstein, A. (1991), "Comments on the Interpretation of Game Theory", *Econometrica* **59**, 909–924. [65, 156]

Rubinstein, A., and A. Wolinsky (1985), "Equilibrium in a Market with Sequential Bargaining", *Econometrica* **53**, 1133–1150. [136, 147]

Rubinstein, A., and A. Wolinsky (1987), "Middlemen", *Quarterly Journal of Economics* **102**, 581–593. [188]

Rubinstein, A., and A. Wolinsky (1990), "Decentralized Trading, Strategic Behavior and the Walrasian Outcome", *Review of Economic Studies* **57**, 63–78. [170, 187, 193, 197]

Sákovics, J. (1993), "Delay in Bargaining Games with Complete Information", *Journal of Economic Theory* **59**, 78–95. [66]

Samuelson, L. (1992), "Disagreement in Markets with Matching and Bargaining", *Review of Economic Studies* **59**, 177–185. [148]

Schelling, T. C. (1960), *The Strategy of Conflict*, Harvard University Press. [6]

Selten, R. (1965), "Speiltheoreticshe Behandlung eines Oligopolmodells mit Nachfrageträgheit", *Zeitschrift für die gesamte Staatswissenschaft* **121**, 301–324. [43]

Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory* **4**, 25–55. [77, 78]

Selten, R. (1981), "A Noncooperative Model of Characteristic-Function Bargaining", pp. 131–151 in R. J. Aumann, J. C. Harsanyi, W. Hildenbrand, M. Maschler, M. A. Perles, J. Rosenmüller, R. Selten, M. Shubik, and G. L. Thompson, *Essays in Game Theory and Mathematical Economics*, Mannheim: Bibliographisches Institut. [67]

Sengupta, A., and K. Sengupta (1988), "Negotiation of Contracts under

Adverse Selection", unpublished paper, University of Western Ontario. [119]

Shaked, A. (1994), "Opting Out: Bazaars versus 'Hi Tech' Markets", *Investigaciones Económicas* **18**, 421–432. [65, 187]

Shaked, A., and J. Sutton (1984a), "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model", *Econometrica* **52**, 1351–1364. [65, 188]

Shaked, A., and J. Sutton (1984b), "The Semi-Walrasian Economy", Discussion Paper 84/98 (Theoretical Economics), Suntory Toyota International Centre for Economics and Related Disciplines, London School of Economics. [65]

Shapley, L. S. (1969), "Utility Comparison and the Theory of Games", pp. 251–263 in G. T. Guilbaud (ed.), *La Décision*, Paris: Editions du Centre National de la Recherche Scientifique. [27]

Shubik, M. (1982), *Game Theory in the Social Sciences*, Cambridge: MIT Press. [25]

Stahl, D. O., II (1990), "Bargaining with durable offers and endogenous timing", *Games and Economic Behavior* **2**, 173–187. [67]

Ståhl, I. (1972), *Bargaining Theory*, Stockholm: Economics Research Institute, Stockholm School of Economics. [65]

Ståhl, I. (1977), "An N-Person Bargaining Game in the Extensive Form", pp. 156–172 in R. Henn and O. Moeschlin (eds.), *Mathematical Economics and Game Theory*, Berlin: Springer-Verlag. [65]

Ståhl, I. (1988), "A Comparison Between the Rubinstein and Ståhl Bargaining Models", Research Paper 6347, Economics Research Institute, Stockholm School of Economics. [66]

Sutton, J. (1986), "Non-Cooperative Bargaining Theory: An Introduction", *Review of Economic Studies* **53**, 709–724. [65]

Thomson, W. (forthcoming), *Bargaining Theory: The Axiomatic Approach*, Boston: Academic Press. [27]

Van Damme, E. (1987), *Stability and Perfection of Nash Equilibria*, Berlin: Springer-Verlag. [x, 78, 90, 112]

van Damme, E., R. Selten, and E. Winter (1990), "Alternating Bid Bargaining with a Smallest Money Unit", *Games and Economic Behavior* **2**, 188–201. [66]

Vincent, D. R. (1989), "Bargaining with Common Values", *Journal of Economic Theory* **48**, 47–62. [119]

Wagner, R. H. (1984), "A Noncooperative Solution to a Two-Person Bargaining Game", unpublished paper. [67]

Wilson, R. (1984), "Notes on Market Games with Complete Information", unpublished paper, Graduate School of Business, Stanford Univer-

sity. [187]

Wilson, R. (1987), "Game-Theoretic Analyses of Trading Processes", pp. 33–70 in T. F. Bewley (ed.), *Advances in Economic Theory*, Cambridge University Press. [6]

Wolinsky, A. (1987), "Matching, Search, and Bargaining", *Journal of Economic Theory* **42**, 311–333. [67, 148]

Wolinsky, A. (1988), "Dynamic Markets with Competitive Bidding", *Review of Economic Studies* **55**, 71–84. [148]

Wolinsky, A. (1990), "Information Revelation in a Market with Pairwise Meetings", *Econometrica* **58**, 1–23. [148]

Zeuthen, F. (1930), *Problems of Monopoly and Economic Warfare*, London: George Routledge and Sons. [16]

Index

The most important entries are indicated by italicized page numbers.